Three Essays on the Economics of Contracts in Labor and Corporate Debt Market

Ding Yuan

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2018

ABSTRACT

Three Essays on the Economics of Contracts in Labor and Corporate Debt Market

Ding Yuan

I study the roles contracts play in the labor and corporate debt market.

Chapter 1 studies wage contracts and their roles in workers' employment and wage dynamics, as well as the implications on income inequality. I develop an on-the-job search model that allows for different types of wage contracts. Using indirect inference method, I am able to estimate the structural model and evaluate the impact of different productivity elements, including firm productivity, returns to routine task and individual effort. The model is able to capture key measures on worker's labor market mobility, wage growth and distribution. It also allows me to evaluate the implications of productivity change on income inequality through counterfactual analysis. I show that these productivity elements have different implications on income inequality, and the use of performance based wage contract is an important channel for income polarization at the top percentiles.

Chapter 2 studies the effect of overtime pay on workers' working schedule and income. How overtime pay regulations affect the labor market is a controversial yet relatively understudied topic. In this paper, I study the effect of the revision to statutory overtime pay in 2004 on worker's income and hours of work. Using monthly panel data on workers' working hours and income that covers the period of rule change, I find evidence that for workers who gained statutory overtime pay coverage under the new rule, hours and income increased. I also find spillover effects on overtime pay premium and overtime schedule for workers who are not directly affected by the rule change. My results suggest that the standard competitive model does not capture well the labor market for overtime work, and government regulations could reduce labor market frictions.

Chapter 3 studies debt covenant violations and their effects on corporate innovation. Exploiting the state of debt contract covenant violation and the institutional feature that creditors obtain increased control right of the firm, the paper examines the effect of increased creditor governance well before the state of bankruptcy on corporate innovation. Consistent with the view that increased creditor monitoring has disciplining effect on the managers, I find no significant change in the R&D spending, significant but model decrease in the total patent counts two years forward as well as significant and large positive impact on the citation counts of the patents. The results demonstrate that increased creditor governance is overall beneficial to firm innovation.

Contents

List of Tables

List of Figures

Acknowledgement

I am deeply indebted to the awesome mentorship of Prof. W. Bentley MacLeod, whose teaching, support, and encouragement made this five-year journey truly rewarding for me. I am grateful for the teaching and advices from Prof. Jushan Bai, Prof. Suresh Naidu, Prof. Tobias Salz and Prof. Stephen Sun Teng. I am grateful for the financial support from the Economics Department and the Graduate School of Arts and Science, which made this possible. I thank Amy Devine and Shane Bordeau for their excellent administrate support. My dissertation has also benefitted from comments from other members of the Columbia Economics community: Prof. Pierre-Andre Chiappori, Prof. Wojciech Kopczuk, Prof. Brendan O'Flaherty, Prof. Miikka Rokkanen, and various other seminar participants.

I am also thankful to my friends and fellow Ph.D classmates for their friendship, support, and collaboration. In particular, Pierre Poth, Anna Lin, Yang Jiao, Tong Geng, Tuo Chen, Sakai Ando, Ye Zhang, Ashna Arora, Daniel Rappoport, Xing Xia, Lin Tian, Adrian Tang.

To the lives and souls I come across in this city,

for teaching me what it means to be a human being.

Chapter 1

Wage Contract, Employment Dynamics, and Income Inequality

1.1 Introduction

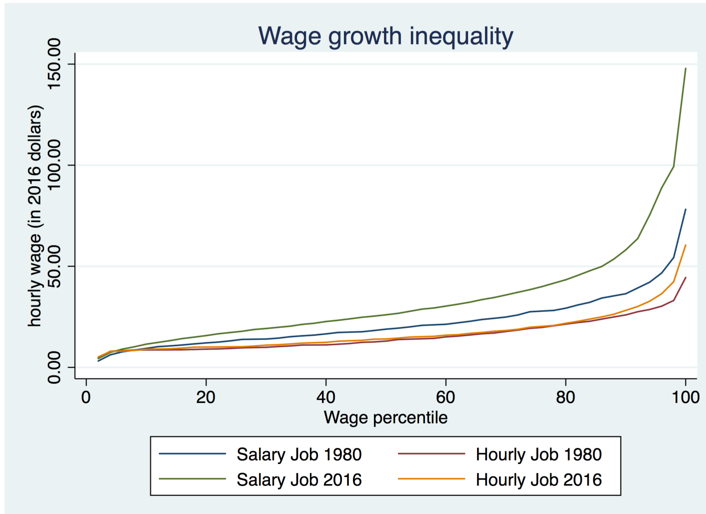
There has been a staggering increase in income inequality over the past three decades both in US and the rest of the world. The rapid concentration of wealth at the top one percentile is empirically well-documented in Piketty and Saez (2003),Piketty and Saez (2006) and Atkinson and Piketty (2007). Economists have put forward many explanations for the increase in income inequality, including skill-biased technological change and the outsourcing of manufacturing jobs. In this paper, we study the wage and employment dynamics under performance pay and fixed wage contracts, and use a structural model to quantify the contribution of difference factors to income inequality. We find that performance pay contracts relative to fixed contracts are associated with higher income, longer working hours, faster income growth, and more stable employment relationship. Using a dynamic job search model, we were able to estimate the distribution of various productivity factors in an employment relationship, including firm productivity, returns to routine task and individual effort, and evaluate the impact of performance pay based employment relationship on income inequality. Our counterfactual analysis shows that change in individual performance and skill based productivity has a much larger impact on income distribution at the top percentile than change in productivity based on routine tasks, and the use of performance pay based wage contracts is an important channel to translate productivity growth into income polarization at the top percentiles.

Over the last 40 years there have been a number of structural changes in US labor markets. Starting from the 1970s, the presence and importance of labor unions diminished steadily. At the same time, the use of performance-based labor compensation had become more prevalent in the labor market, so had income inequality. (Lemieux, Macleod and Parent (2009)) In recent years, the burgeoning mobile information technology has brought the so-called "gig-economy" (in which employment relationships are formed for specific "on-demand" tasks without long-term commitment) to many industries.[1] These issues necessitates the efforts to understand the labor market through the lens of employment relationship.

Labor contract plays a very important role in shaping the relationship between workers and their employers. How wage is determined is one of the key terms in the labor contract. We can roughly put wage compensation into two categories: fixed wage and performance pay. Under a fixed wage job, workers will receive pre-agreed amount of labor compensation at specified frequency, whereas in performance pay, labor compensation will have a variable pay component that depends on the eventual output or the quality of work, for example piece rate, commission, bonus or tips are all different forms of performance pay. Heuristically, employment relationship with performance pay contracts should be relatively more stable under demand or productivity shocks: workers are incentivized to exert more effort to achieve higher level of production, and they will share part of the risk with employers by taking a cut in the compensation when there is a negative shock. There are quite a number of paper on this issue. Azariadis (1975) and Beaudry and DiNardo (1991) show that fixed wage contract is a way to provide insurance for risk-averse workers. On the other hand, Rosen (1985) argues that if wage contracts are designed to optimally share risk with workers, then this implies workers who are laid-off are not necessarily worse off, which seems inconsistent with the evidence. On the empirical side, Lemieux, Macleod and Parent (2009) shows that labor compensation are more closely tied to worker attributes in performance pay jobs, and that provides a channel for change in return to skills to translate into higher wage inequality. Lemieux, Macleod and Parent (2009) further shows that wages and working hours are more responsive to local labor market shocks for performance pay jobs, but less responsive when it comes to employment. Figure 1.1. shows the wage growth disparity over the past 4 decades between jobs under different types of wage structures. While hourly-paid jobs hardly see much real income growth across the entire income distribution during the period, salary

[1] For example, Uber in the transportation service industry, AirBnb in the hospitality industry, Handy in the domestic-service industry

Figure 1.1: Wage Growth Disparity



Data Source: Current Population Survey (CPS)

Wage is hourly wage for salary and hourly paid jobs, in 2016 dollars. In CPS survey, interviewees are asked if they are paid by the hour. Hourly job is defined as jobs where income is hour-based.

jobs have seen a lopsided growth concentrated at the top. For salary jobs, wage growth is much larger in scale for income at 90th percentile and above. This shows that on top of the wealth concentration across income distributions, income inequality also has a group dimension: certain groups have seen increased income disparity while other groups less so. Even though the group distinction between salary and hourly jobs is not exactly cut along the performance pay - fixed wage line, majority of the hourly paid jobs are non-performance pay based, and performance pay jobs tend to be concentrated on the top percentiles of salary jobs.

In this paper, we seek to better understand the impact of labor contracts on workers' income and employment dynamics, and the contributions of different labor market factors to labor income inequality. More specifically, we focus on the following: 1). Document the differences in labor market mobility and wage dynamics for job spells under different wage contracts; 2). Adapt a dynamic equilibrium job search modeling framework to incorporate both performance pay and fixed contract to describe individual employment and income dynamics; 3). Estimate the structural parameters of the model to study the differences of productivity factors under different contract group; 4). Evaluate the contribution of different productivity factors to the cross-sectional income distribution and income growth inequality.

Using the PSID, we are able to group job spells into fixed wage or performance pay. We mainly focus on two sets of dynamics: individual labor market mobility and wage dynamics. For labor market mobility, we look at the transition rate of the entire worker population in the sample (job to job transition and job to unemployment transition), conditional on their tenure level. We also non-parametrically estimate the survival function for the two types of transitions. For wage dynamics, we use the Mincer wage and wage growth equations to estimate how worker's earnings evolve over time. We performed the estimations on all the education and wage contracts groups. Through the analysis, we have two major findings: 1). Within the same education group, there are very big differences in labor market mobility, workers under performance pay contracts are much less likely to switch jobs or get fired than those under fixed wage contracts. Meanwhile, within each contract group, the gap in labor market mobility is much less significant. 2). Returns to tenure as estimated from the Mincer Wage Equation are higher for performance pay jobs in low and medium education groups[2], but there's a reversal in the high education group. As the main subject of interest in our work is the cross contract group differences, the economic significance of our empirical result is central to our thesis. One of the pitfalls of our empirical approach is that the the choice of wage contract might be systematically correlated with job and worker characteristics, if that's the case, then a cross group comparison could tell us very little about the effect of contract itself. To mitigate this problem, we also study the subsample of individuals who had worked under both type of wage contracts during their career, so that confounding factors related to individual characteristics can be controlled. From the subsample, we also get the same robust empirical result. We also show the income distribution and inequalities between

[2]We group individuals with education years less than 12 as low education group, between 11 to 14 as medium education group, and above 14 as high education group.

different contract groups along a few measures: cross-sectional distribution of income, income growth disparity at difference income percentiles, and income inequality among workers over long term aggregate labor income. We show that there is a relatively big income gap between top earners in performance pay jobs both within contract group, and also compared to fixed wage job group. Performance pay jobs also have much larger income growth rate than fixed wage job at the top income percentile. Furthermore, prior to 1985, worker's long term labor income were comparatively evenly distributed, but after 1985 there starts be an increasing income gap between top percentiles and the rest. In addition, the income gap at top percentile between contract groups has also widened since then.

The main objective of our model is to describe the labor market process for workers under different compensation schemes that can be easily applied to empirical analysis. Our model combines two strands of literature in labor economics to study the dynamics of wage and employment under different employment contract. The first strand is the literature on wage contracts. We mainly adopt the modeling approach in Lemieux, Macleod and Parent (2009) where they use a single period wage model to highlight the differences in performance pay and fixed wage contracts. Wage contract differences are modeled as a function of the worker's effort and its effect on production. We endogenous the choice of wage contract with monitoring cost of effort: it captures the difficulty of enforcing worker's effort, or the measurability of output in order for it to become contractable. The second second strand is the structural labor. We mainly follow the modeling framework in Bagger et al. (2014), in which they developed an equilibrium job search model to describe worker's career dynamics in terms of income and employment.

There are three features of the model worth highlighting: 1). We adopt the random on-the-job-search framework in which workers have certain probably of receiving outside job offers; 2). We add endogenous termination of employment relationship into the model, as a result of firm's and worker's rational choices based on new information, we consider this to be an important extension to many of the existing random on-the-job-search models, in which worker-employer separation is exogenous; 3). We include working hours into the model as an endogenous factor, it affects both worker's utility and production outcome, and is a result of rational choice based on information about productivity shock. This extends the dimensions on which the model is able to describe. In the end, our modeling framework is able to yield a comprehensive description of worker's labor market experience that can be easily applied for empirical analysis.

We use indirect inference to structurally estimate the model using the long panel from PSID on worker's employment and labor income. The employment and earning dynamics for workers are the key aspects of our interests, therefore we rely heavily on labor market mobility and Mincer Wage Equations for estimation. Our model is able to produce good fit for worker groups under both types of wage contracts, and across education groups as well. This validates our model in terms of describing the labor market process. We view this as central to our thesis: it shows that conventional studies using "one size fits all" modeling for labor market data with a mixture of contractural agreements are missing important aspects of the labor marker process, and the estimation results could be potentially biased. This highlights the need to incorporate contractural aspects in the empirical model, which is objective of our work. With our structural model and estimated parameters, we evaluate the effect of different factors on the overall labor market income distribution. Our counterfactual analysis shows that changes in different productivity elements have different implications for income inequality, especially considering wage contracts, returns to effort have a much larger effect on income polarization at the top end, and the use of performance pay contract is an important channel for income polarization.

We consider our work to contribute to three strands of literature. First, we contribute to the literature of income inequality by developing a structural model to evaluate the effects of various labor market factors, and in particular, the role of wage contract. The role of performance pay and their contribution to income inequality has been quite extensively documented and discussed in Lemieux, Macleod and Parent (2009). As they point out, performance pay channels increase in returns to skill into income inequality. Our work extends on that basis, and instead of looking at cross-sectional evidence, we study the dynamics of worker's labor market experience and how that translates into income inequality. Therefore, it allows us to look at worker's income difference over a more extended period of time, and include more labor market factors. In Lemieux, Macleod and Parent (2009), the difference between contracts lies in monitoring cost and is static, by using a dynamic model, we include more factors such as firm level productivity difference, worker's human capital accumulation, and worker's change in income through switching jobs, being dismissed from worker, and renegotiate wage contract using outside offers. By doing this, we were able to construct different counterfactual experiments to evaluate the contribution of different factors to income inequality, for example, the change in productivity in effort, or lowering of monitoring cost, individual productivity heterogeneity, and fluidity of labor market transitions. Our work also brings a new perspective into the research on "skill biased technological change", such as the work in Acemoglu and Autor (2010), Autor and Dorn (2013). Their paper showed the effect of technology development, the automation of routine tasks on the polarization of labor market. By modeling returns to different productivity factors, our approach allows to construct counterfactual examples to evaluate the effect of growth of different type of productivities on income inequality, for example, the increased heterogeneity in returns to working hours (routine tasks) has a much smaller impact on income inequality than the increased heterogeneity in returns to effort (skill based abstract tasks). Finally, our work contributes to the study of development of income inequality over the past few decades by studying the labor market dynamics over an extended period of time in PSID data, such as the work in Kopczuk, Saez and Song (2010), and Piketty and Saez (2003).

Our second contribution is to make extensions to the empirical study of labor contracts by exploring the cross contract group differences in workers' employment and earnings dynamics. Previous work has focused on the cross-sectional and time series properties. Lemieux, Macleod and Parent (2009) is about the returns to observed and unobserved worker characteristics; Lemieux, MacLeod and Parent (2012*a*) studies the labor market adjustment to demand shocks at individual level; while Lemieux, MacLeod and Parent (2012*b*) is concerned on the autocovariance structure of wages and earning. In comparison, our work makes job spell the main subject of interest, and emphasizes labor market mobility as the key aspect of job spell. In doing so, we essentially used the methodology of structural labor literature to study labor contracts. We find strong evidence that labor market mobility are significantly different under the types of wage contracts: workers under performance pay contracts tend to have more stable employment than those under fixed wage contracts, less likely to switch job or get laid-off.[3] This suggests that structural labor work in which the models does not differentiate contractual terms might be missing an important aspect of the model. Our structural estimation results further affirms this point: models allowing for different types of wage compensation performs much better in replicating key features of the data, as in comparison with singular wage contract modeling. In this way our work is trying to bring these two strands of literature together: the empirical evidences on labor contracts points out the need for more flexible contract modeling in structural labor work, and our structural modeling and estimation results provide additional perspective into the impact of wage contracts on worker's labor market experience. Through the structural approach, we are able to quantitatively analyze how wage flexibility affects workers' employment stability. Specifically, the key difference between our work and previous empirical study on labor contracts is that structural model allows us, to some extent, back out the unobserved firm and worker heterogeneity from the data, so that we can control for these factors and use counterfactual analysis to quantify the impact of wage contract itself. This becomes appealing when, on the one hand, the choice of wage contracts is related to a lot of unobserved factors such as the the level of uncertainty faced by the firm or worker, therefore a clean reduced-form identification is extremely hard to achieve; [4] and on the other hand, it's very rare to

[3]As convention in the literature, we divide workers into three groups by their education levels, our results are consistent across all the groups

[4]Ideally, identification needs randomized experiments in which workers are randomly assigned to different

have high quality labor market data that contains both detailed firm-level and worker-level information over extended period of time, which would be extremely helpful for researches on labor market dynamics.

Finally, we add to the structural labor literature a framework that allows for different contractural terms when modeling worker's labor market experience. On the subject of individual earnings dynamics, there is a vast amount of literature that used various stochastic models to decompose earnings data, to name a few notable works: Abowd and Card (1989); Meghir and Pistaferri (2004); Guiso, Pistaferri and Schivardi (2005); Gottschalk and Moffitt (2009). In recent years, there are quite a number of works that used Indirect Inference Method to structurally estimate dynamic models on individual's income and employment, notably Altonji, Smith and Vidangos (2013) and Bagger et al. (2014). Our model is built on the framework in Bagger et al. (2014), and extends it by incorporating productivity of effort, working hours, and different wage contracts. In particular, we model the endogenous separation process and different types of contractural commitments to wage compensation, which, to the best of our knowledge, is new to the literature. In our model, wages and working hours are consequences of contractural bargaining between workers and employers; workers change jobs through on-the-job search; employment relationship terminates due to ex-ante commitment and ex-post negative productivity shocks. Our work shows that the modeling framework in Bagger et al. (2014)Bagger, et al (2014) can be easily extended to study a wide range of issues in labor economics.

The rest of the paper is organized as follows. In Section 2, we discuss the empirical results on worker's employment and wage dynamics, by their education and contract groups. In Section 3, we present a modeling framework that describes worker's labor market experience over extended period of time, allowing for a number of heterogeneities and different wage compensation scheme. In Section 4, we describes our estimation methodology and shows

wage compensation schemes and all the other labor market factor such as firm productivity, demand shocks are kept the same for workers in different experiment groups

Count	Edu 7-11	Edu 12-14	Edu 15-20
(Worker)			
Fixed Wage	932	2922	979
Performance Pay	193	1009	508
Total	981	3232	1133
(Job Spell)			
Fixed Wage	2678	8098	2525
Performance Pay	212	1155	657
Total	2890	9253	3182

Table 1.1: Worker and Job Spell Sample Count across Groups

the structural estimation results. In Section 5, we evaluate the contributions of different productivity factors to income distribution and wage growth inequality.

1.2 Wage and Employment Dynamics

1.2.1 Data

We use the same PSID dataset as in Altonji, Smith and Vidangos (2013) and Lemieux, Macleod and Parent (2009). Worker sample consists of those who were ever household head during the period from 1975 to 1996. We future restrict the sample into those who were not self-employed, have at least four consecutive periods that were in the labor force, and had at least a job spell longer than 3 periods. In total we end up with roughly 5000 individuals, 60000 person-year and 15000 job spells.[5] Following convention in the literature, we divide workers into different groups, by their education level: high (15 to 20 years), medium (12 to 14 years) and low (7 to 11 years).

[5]We are very careful in following standard producers in the literature (a good example is Altonji, et al (2013)) to process the original PSID data, as it is well-documented that variables such as "tenure" are not well documented and needs different methodology to back out and clean. The STATA do file for the data cleaning process is available upon request.

Throughout the period, workers in the PSID were asked questions relevant to their wage compensation, such as how were they being paid: in piece rate, commission, bonus, salary, tips, or commissions, and whether they've received performance pay in addition to regular salary. We adopt the same method in Lemieux, Macleod and Parent (2009) to identify performance pay jobs: job spells that are specifically recorded as performance pay jobs or the worker had ever received compensation in the form of piece rate, bonus, tips or commissions are identified as performance pay jobs (excluding overtime pay). Table 1.1 shows the sample count on worker and job spell across groups. The ratio of worker and job spell under performance pay contracts ranges from about one-tenth to one-third, and increases with education level. Middle education group has the largest sample count.

1.2.2 Labor Market Mobility

Employment duration is a very important aspect of job spells. Employment relationship that lasts for a long period of time is considered to be stable, while jobs that are volatile usually end with workers becoming unemployed or switching to a new job very soon. There are different ways to look at labor market mobility. First, we consider on average how stable is an employment relationship conditional on workers' tenure. Figure 1.1 shows the ratio of job spells in which the employment relationship continues for each given level of tenure. The ratio of job continuation is consistently higher for jobs under performance pay contracts, across all the tenure levels and education groups. Figure 1.2 and Figure 1.3 display the ratio of jobs ended with workers switching to another employer or being laid off, respectively.

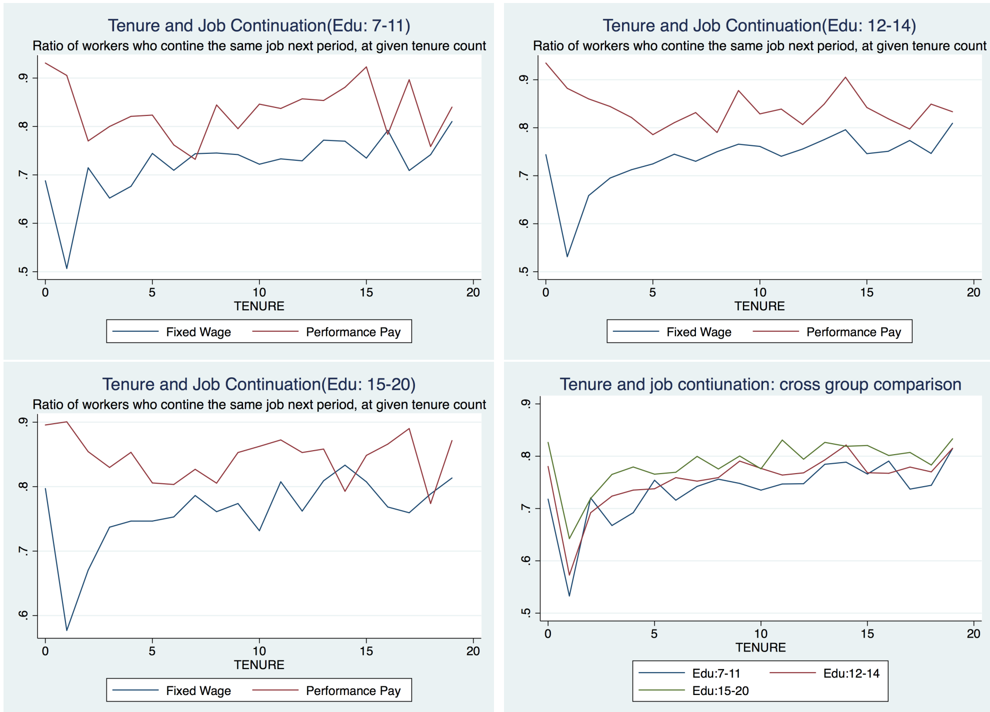


Figure 1.1: Tenure effect on job continuation for different education groups and wage contracts

We see a very distinctive pattern on these measures through the comparison of different wage contracts: jobs under performance pay are much more stable than fixed wage jobs: conditional on years of tenure, less fraction of the employment relationship end with job switching or separation, and a larger portion of employment relationship continue to the next period.[6] It is noticeable that there tend to be a dip/bump on the graphs at the first year tenure point. One of the reasons is that there is a significant number of cases in which the employment relationship only lasted for a year, probably due to mismatching and the termination of employment ensued. There is also the fact that a fraction of the tenure levels and job spells are imputed using standard producers, hence measurement errors tend to have a larger impact on the first few years of each job spell. This does not undermine the significance of the results because of the consistent pattern across groups and going beyond

[6]During the period of our data coverage, PSID provided information annually, therefore our unite of time period is in year.

the first few noisy years of tenure measure.

Figure 1.2: Tenure effect on job switching for different education groups and wage contracts

Figure 1.3: Tenure effect on job separation for different education groups and wage contracts

To take the above result a step further, we look at the employment-to-employment transition (workers switching jobs without experiencing unemployment, hereafter EE transition) and employment-to-unemployment transition (hereafter EU transition) by nonparametrically estimating their corresponding survivor function using the Kaplan Meier procedure. It enables to deal with the problem of right-censoring commonly encountered in estimating survivor functions.[7] This is also the estimates used in Bagger et al. (2014). The K-M estimates for EE transition is given by:

$$\hat{S}_{EE}(\tau) = \Pi_{s=1}^{\tau} \frac{|\mathcal{P}_{EE}(s)| - |\mathcal{D}_{EE}(s)|}{\mathcal{P}_{EE}(s)}$$

$\mathcal{P}_{EE}(\tau)$ is the set of spells at risk of ending in a job-to-job transition at duration τ. $\mathcal{D}_{EE}(\tau)$ is the set of spells that do end in a job-to-job transition at duration τ. The EU transition

[7]As a job spell terminates and registers as a realized case, it also drops out of the entire sample space.

estimates are defined analogously.

Table 1.2 shows the estimates across education groups and wage contract types. The results are in first-differences of the survivor function, which represent the transition propensity across each period. Figure 1.4 and 1.5 plot the first-differenced estimates on EE transition rate across contracts and across education groups, respectively. Figure 1.6 and \ref{fig:EU_cross_edu} compares estimates on EU transition rate. These survival function estimates reassures the previous results on labor market transition ratios conditional on tenure.

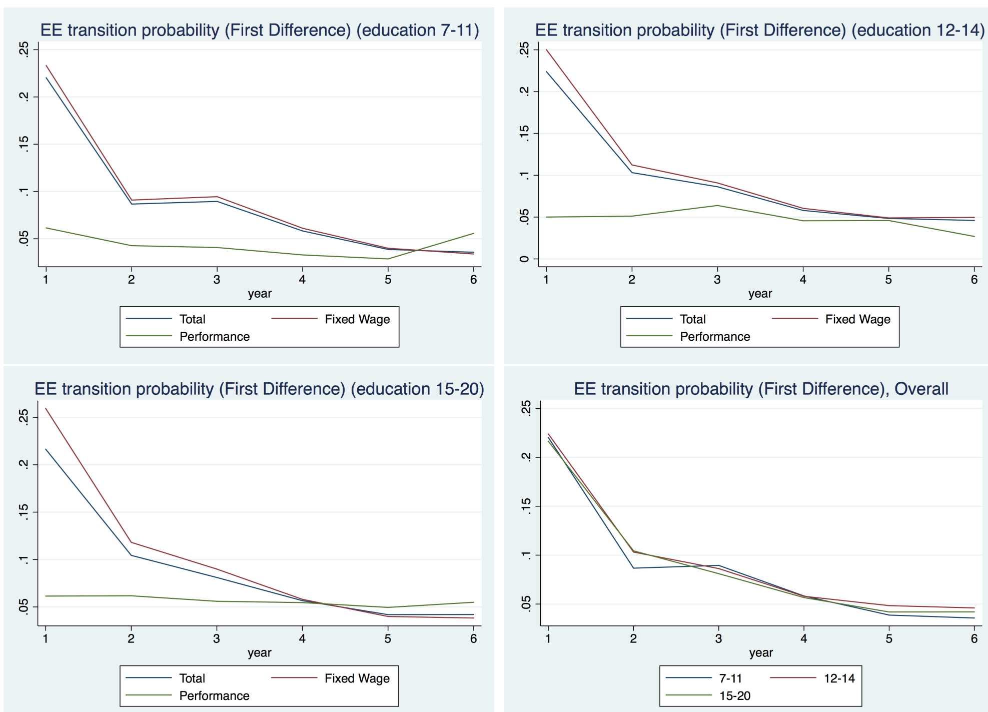


Figure 1.4: EE transition: Within Education Group Comparison

From the cross-contract comparison in Figure 1.4 and 1.6 we find that workers under performance pay contracts are subject to less likelihood of termination of employment, either through changing employers or becoming unemployed. The estimates on first-differenced survival rate in EE or EU transitions are often several times larger under fixed wage contract. For example, in the middle education group (12-14 years), the first and second period difference in EE transition survival rate are 0.1123 and 0.0511 for fixed wage and performance pay groups; likewise for EU transition, the measures are 0.0541 and 0.0089 for these two groups, an even more significant difference. A larger magnitude of estimate means higher likelihood of transition across periods. From the cross-contract within-group figures we see that the disparity in labor market mobility is bigger for EU transitions than EE transitions. EE transition difference gap closes quickly after the first few periods whereas EU transition gap are relatively more persistent during a job spell. We consider this to be an interesting result as it not only shows the systematic differences in the stability of employment relationship, but also drops hints on the underlying mechanisms that drives the differences: one plausible explanation is that the EE transition is highly related to worker-job matching quality and workers' job search intensity, hence a group difference in these aspects could lead to the survival rate gap initially, but as workers settle into their jobs the employment relationship becomes more robust, as a result the group gap is closing; The persistent contract group difference in EU transition rates might be pointing to the constant effect of wage compensation method on firm's laying off decisions.

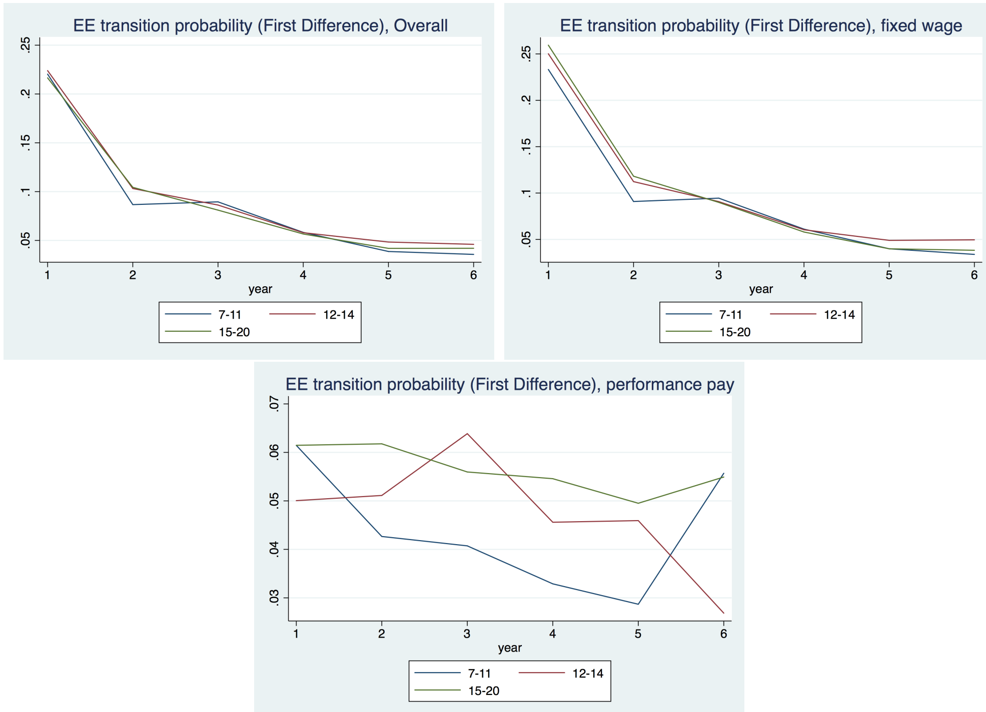


Figure 1.5: EE transition: Cross Education Group Comparison

Figure 1.5 and 1.7 exhibit the comparison across different education groups. For EE transition, there isn't strong evidence for group-wise differences. For job spells under fixed wage, the EE transition rate appears to be very similar across the three education groups; for job spells under performance pay contracts, the EE transition rates do not have a systematic group difference. For EU transition, in the fixed wage category, higher education group has smaller transition rate than lower education group, suggesting that employment relationship tend to be more stable for high-skilled worker; meanwhile the transition rates are very close across education groups among the performance pay category.

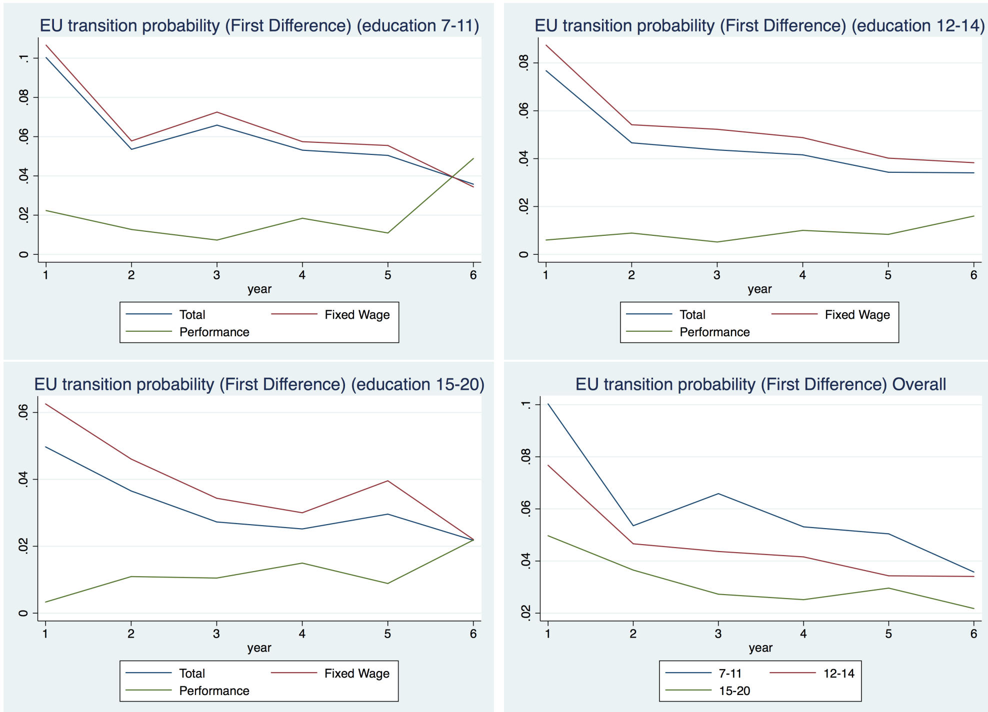


Figure 1.6: EU transition: Within Education Group Comparison

Throughout the statistics and plots, one result stands out: employment relationship under performance pay contracts is more stable than those under fixed wage contracts in terms of workers switching jobs or being laid off. This is central to one of our key arguments: the flexibility of wage compensation allows firms and workers to better adjust to uncertainties, through mechanisms such as ex-post risk-sharing and optimal production, as a result the worker-job matching is more stable and more resilient to employment risks. As such, having these simple comparison of estimates is not an identification strategy for the causal effects of wage contracts. In particular, there are characteristics of jobs and workers that are systematically different under the two types of wage contracts. For example, performance pay jobs tend to be more managerial (in the form of bonus payment), require higher skills and higher education level. Confounding factors like these mean that the cross-contract differences in labor market mobility should not be attributed to wage contract itself. However, there is still very strong evidence on the role of wage compensation. First of all, by grouping workers by their education level, the effects of confounding factors such as worker skill and seniority in the career ladder are to some extend controlled, and yet within each education group, the differences between contracts are still significant and consistent. Moreover, the cross-education group differences have a quite different pattern from the cross-contract differences: EE transition differences in cross-education comparison is very little whereas significant and diminishing with job tenure in cross-contract comparison; the magnitude of EU transition differences in cross-contract comparison is much larger and more persistent than in cross-education comparison. To the extend that high versus low education level and performance pay versus fixed wage have similar group-wise differences in human capital and rank in the organization, the fact that cross-education comparison does not match cross-contract comparison is strong evidence that wage contract itself plays a significant role in labor market mobility. To affirm this point, we also look at the sub sample that consists of workers who had ever worked under both types of wage contracts throughout their careers. For the sub sample, we end up with 1252 workers, 1512 job spells under performance pay contracts and 3117 job spells under fixed wage contracts. Figure 1.8 shows the comparison for both EE and EU transition. The difference between contract groups are essentially the same as in the full sample: workers under performance pay contracts have less propensity to switch jobs, but the difference closes up in a few years; workers under fixed wage contracts are more likely to become unemployed and the gap between contract groups is consistent throughout the job spell. This sub-sample enables us to control for unobserved worker characteristics that are systematically related to wage compensation methods since the workers in the two wage contract groups are exactly the same. All in all, we conclude that there is significant differences in labor market mobility across wage contract groups, and also there's strong evidence that wage compensation method might play a role in it.

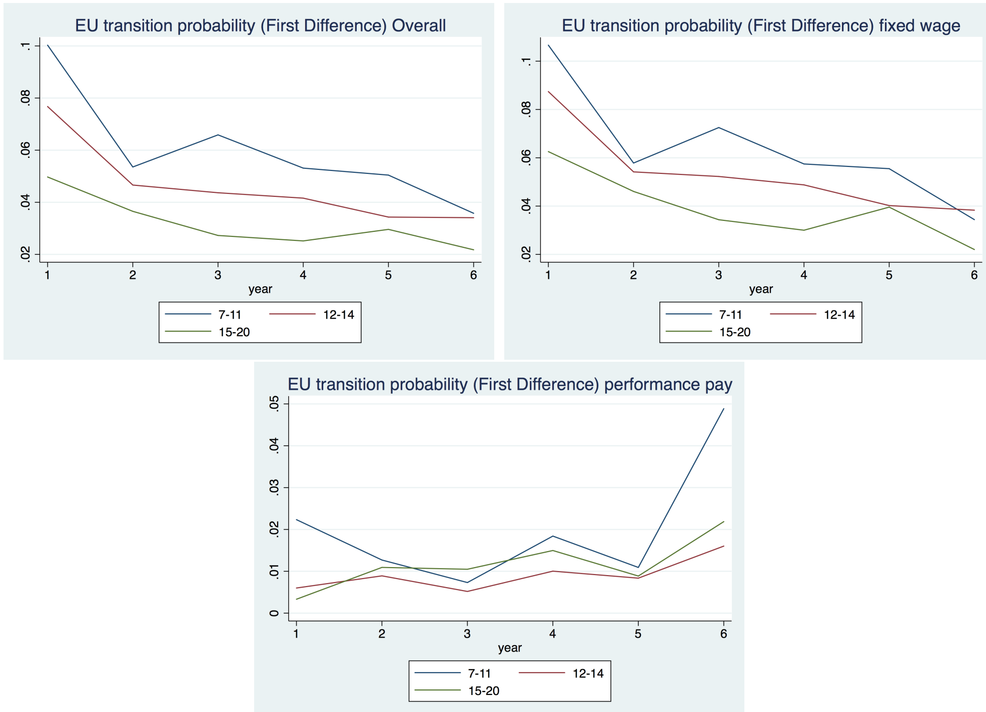


Figure 1.7: EU transition: Cross Education Group Comparison

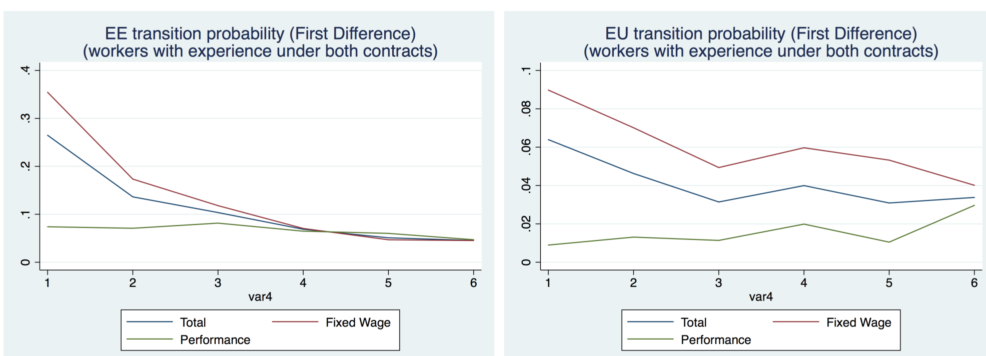


Figure 1.8: EE and EU transition: Sub-Sample Comparison

1.2.3 Working Hours and Wage Dispersion

In addition to labor market mobility, we are also very interested in workers' wages and hours of work. Fixed wage contracts are less flexible in compensations than performance pay contracts. Therefore, the within job spell variations in wages and hours would reflect that on the intensive margin. Under a negative shock, a firm could choose to layoff the worker if her wages and hours are rigid. Alternatively, the firm could choose to pay less and demand less working hours if the wage contract allows for ex-post adjustment.

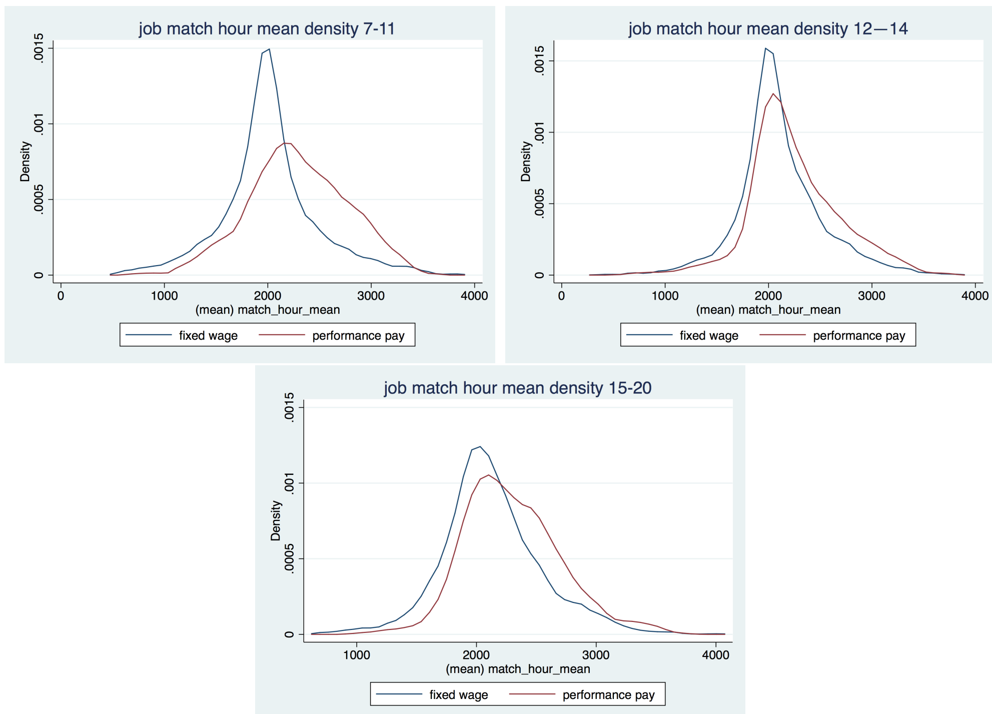


Figure 1.9: Within job spell average annual working hours distribution

Consequently, we should expect to see that within a job spell, wages and working hours tend to have larger variations for performance pay jobs. Table 1.5 reports the summary statistics on the wages and hours of work in the data. On average, performance pay jobs have longer working hours and higher wages. The group averages on annual working hours (fixed wage jobs versus performance pay jobs) are 2018 v.s. 2292, 2104 v.s. 2246, 2013 v.s. 2302 for low, medium and high education groups, respectively. The comparisons for wages are similar.

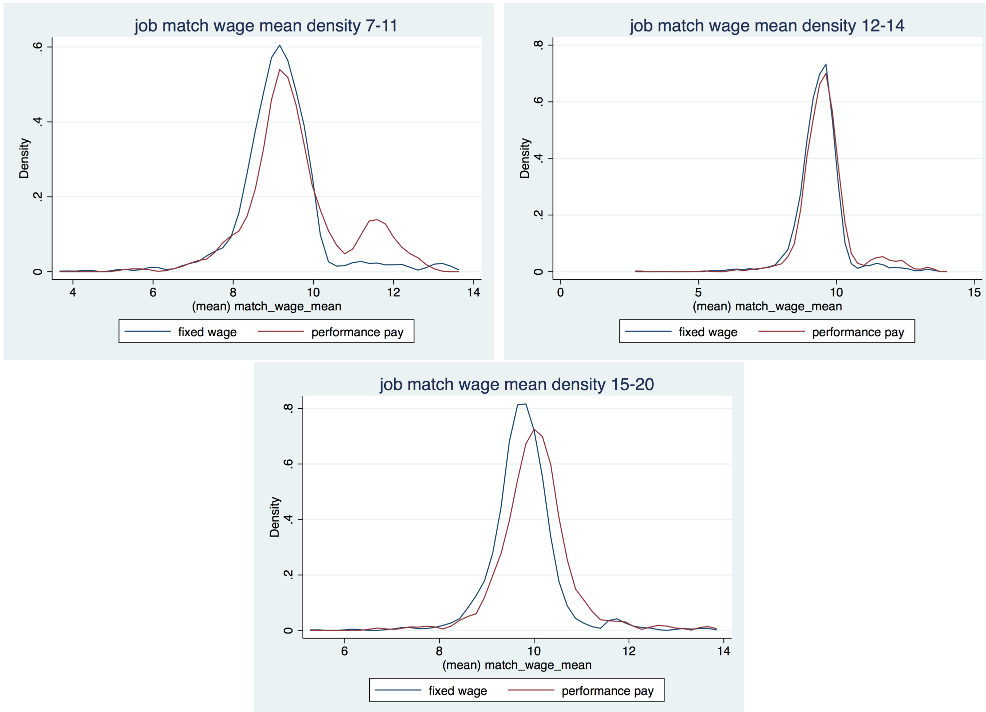


Figure 1.10: Within job spell average log wage distribution

Performance pay jobs also tend to have larger variations in wages and working hours. For each job spell, we measure the variations (in standard deviation) of wages and working hours, then we look at how the statistics are distributed across the entire population in the data. The average level of wage (in logs) variation within each job spell is considerably larger in performance pay jobs. Figure 1.9 and Figure 1.10 plot the density graphs for within job spell average level of annual working hours and wages. Figure 1.11 and Figure 1.12 plot the density graphs for within job spell variations of hours and wages. The differences between the two wage contract groups are evident from these figures. Workers with performance pay jobs tend to work longer hours and receive higher compensations. At the same time, their hours and wages are more volatile.

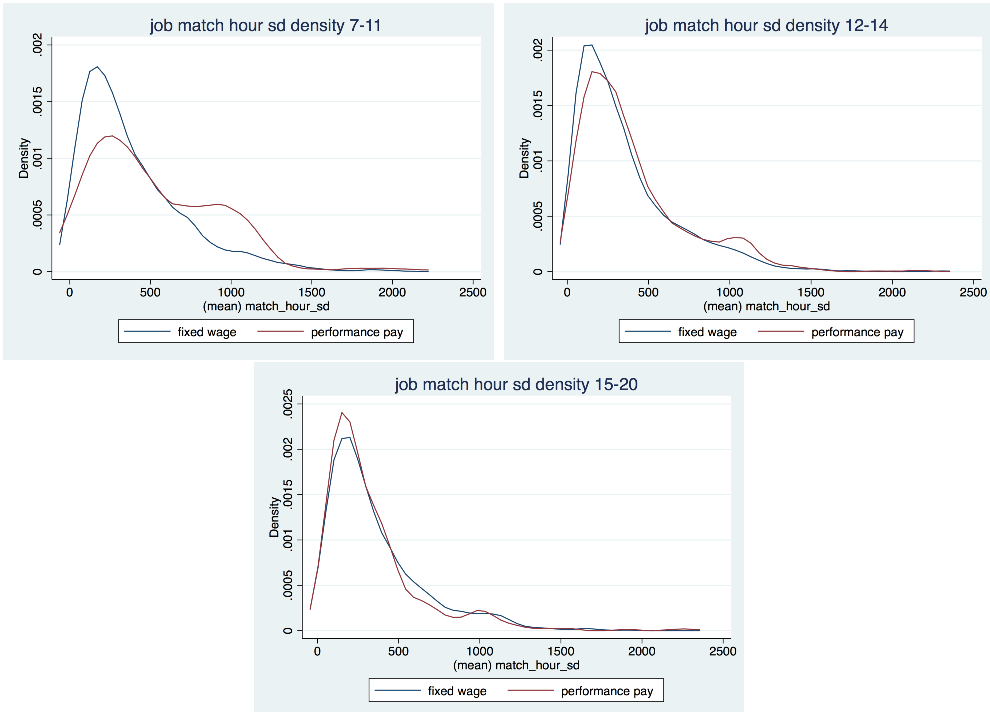


Figure 1.11: Within job spell variation of annual working hours: distribution

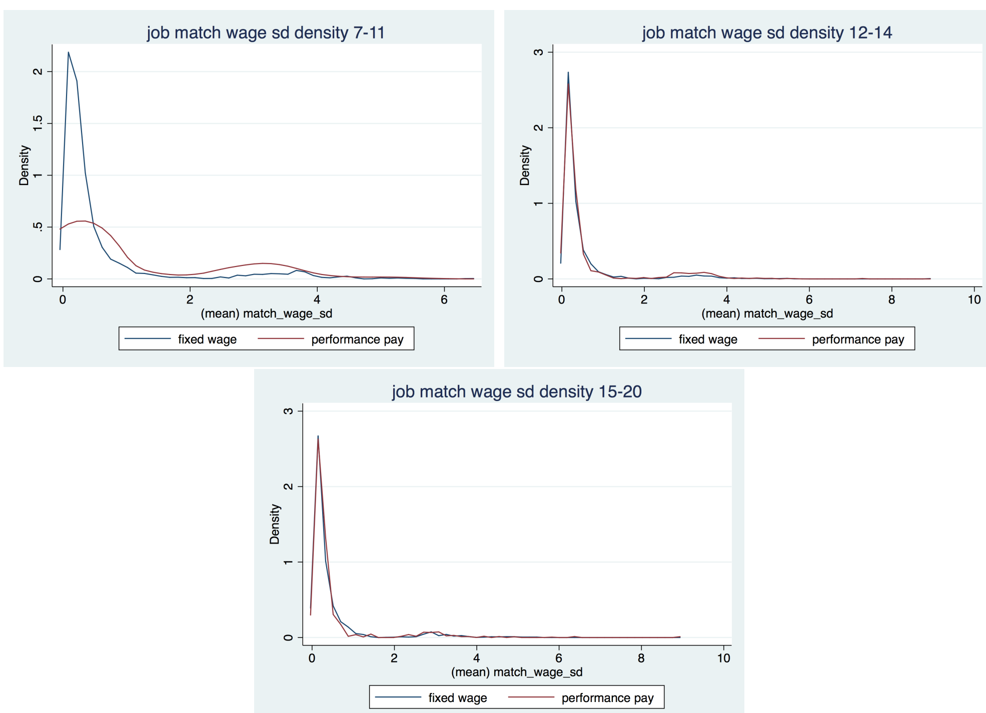


Figure 1.12: Within job spell variation of log wage: distribution

1.2.4 Mincer Wage Equations

We estimate the returns to tenure and potential experience using Mincer wage regression as the following:

$$w_{it} = \sum_{k=1}^{3} \xi_{1k} t_{it}^{k} + \sum_{k=1}^{3} \xi_{2k} s_{it}^{k} + \psi_i + \varepsilon_{it}$$

Where i is job match, w_{it} is log wage for job match i in period t. t_{it} and s_{it} are tenure and potential experience respectively, ψ_i is job match fixed effect. In the regression specification, we use polynomials up to third order to model returns to tenure and potential experience. Table 1.3 reports the estimates across groups. Comparing the estimated numbers within each education group, the wage dynamics are quite different for different contract groups. For example, the linear term coefficient for returns to tenure ξ_{11} are 0.09955 v.s. 0.1820 for fixed wage workers v.s. performance pay workers in low education group. The estimates comparisons for medium and high education group are 0.0744 v.s. 0.1346 and 0.0652 v.s. 0.0434 respectively. Meanwhile, within the same wage contract group, the wage returns are relatively stable across education groups, especially for fixed wage contracts, the fact that we have more samples in that group than for performance pay jobs might have contributed to that. Figure 1.1 plots the wage-tenure profile for each education group. The return profiles are quite different between contract groups but is not consistent across education levels: returns to tenure for fixed wage jobs are lower in lower and medium education groups, but the relationship is reversed in high education group.

1.2.5 Wage Contracts and Income Inequality

The previous sections show the differences of the wage and employment dynamics under different wage contracts. In this section, we show the income distributions and inequalities. Differences in labor market mobilities for workers translate into income inequalities mainly through two channels. First, a more robust employment relationship implies a more steady income stream for the worker. Employees who voluntarily leave their jobs or are dismissed by the employers usually suffer a huge income loss for a prolonged period of time. Second, switching between jobs is a common way for workers to get higher salaries, due to better productivity matching or negotiations for better contract terms. Therefore, workers with more stable employment relationships and more outside job options tend to earn more. The dispersions and fluctuations of working hours are also related to the cross-sectional differences in worker's labor income, since it translates into productivities and workers usually get compensated more for longer working hours. The differences in the growth of wage over a worker's tenure will also translate into income inequality, since the accumulation of working experience and job tenure does not translate into income increase at the same rate.

We illustrate the contract group differences in income distributions by three measures. First, we measure the cross-sectional distributions of income by education and wage contract groups. Second, we measure the cross-sectional differences in income growth at different income levels. Finally, we look at the distributions of aggregate income over a long period of time for workers under different wage contracts.

Figure 1.3 shows the income distributions by education and wage contract groups. We measure the average annual labor income of each job spell in the panel data, and plot the cross-sectional distributions of the average job spell income at each percentile. The first graph plots the comparison between wage contract groups. Under performance pay contracts, there is a large gap of income between those above the 90th percentile and the rest, as shown by the sudden steep increase in the slop of the income distribution line. In comparison, jobs under fixed wage contracts are distributed more evenly on income, as the distribution graph is relatively flat in the top percentile range. The remaining three graphs plot the distributions by education groups. As illustrated, the income gap between the top earners and the rest under performance pay jobs is larger for low and medium education groups, whereas the income distributions of high education groups are relatively flat both within and cross the two contract groups.

Figure 1.4 shows the wage growth rates at different income percentiles. We measure the income growth rate within each job spell as the median of the rolling 2-year annual labor income growth rates. We plot the average wage growth rate of each job spell at their corresponding income percentile, as well as a fitting line that represents the distribution. The first graph shows the overall contract group differences: wage growth rates at the two contract groups are relatively similar below the 80th percentile of income distribution, but there is a large difference at the top percentile: high income earners (above the 80th income percentile) with performance pay jobs have much larger income growth rates than the fixed wage jobs at similar income percentiles. This shows that not only do upper income percentile jobs under performance pay have disproportionally high incomes than the rest, they also have much larger income growth rates, further increasing the cross-sectional income inequality. The remaining three graphs plot the distributions by education group. Similar to Figure 1.3, the income growth rate gap between top earners under performance pay and the rest is larger for low and medium education groups, whereas the income distributions of high education groups are relatively flat both within and cross contract groups.

Figure 1.5 plots the long term labor income distributions by different contract groups. On top of annual labor income distributions and the corresponding differences in income growth, we are also interested in the labor income inequality over an extended period of time. To that end, we look at workers' aggregate labor incomes over a five-year time horizon. We include three difference cases. 1). Starting with the initial period, the total labor income of each worker over the next five years, regardless if she becomes unemployed or switched jobs over the period; 2). Starting with the initial period, the total labor income of each worker over the next five years, not including cases where the worker becomes unemployed, but allowing for switching to new jobs; 3). Starting with the initial period, the total labor income of each worker over the next five years, only including cases where the worker is continuously employed by the same employer. On the left side of the figure, we only plot cases where the initial period is before 1985; on the right side of the figure, we include all the cases. A comparison between the three graphs on the left and the three on the right shows the trend in income distribution: prior to 1985, workers' long term labor incomes were relatively evenly distributed, but after 1985, there was an increasing income gap between the top percentiles and the rest. In addition, the income gap at the top percentiles between the two contract groups has also widened since. The comparison between the graphs on the left also shows the impact of labor market mobility on contract group income gap at the top percentiles: the income gap at the top percentiles between performance pay and fixed wage jobs is larger when we include workers' transitions into new jobs in comparison to the 5-year continuous employment group. This shows that job transitions play an important role in increasing the income gap between contract groups at the top percentile. Workers on high-income performance pay jobs gain even higher long-term labor income by transitioning into new jobs, compared with high income fixed wage jobs. Likewise, comparing the top chart with the middle chart shows that employment stability also plays an important role in the income gap at the top percentiles. As shown in the labor market mobility estimates, performance pay jobs are more stable in terms of involuntary separation, this means workers under performance pay jobs are more likely to maintain a stable income stream over a long period of time.

1.3 Model

Our model combines two strands of literature in labor economics to study the dynamics of wage and employment under different wage contracts. The first strand of literature is on wage contract. We mainly adapt the modeling approach in Lemieux, Macleod and Parent (2009), where a single period wage model is used to highlight the differences between performance pay and fixed wage contracts. The second strand of literature is on structural labor search models. We follow the modeling framework in Bagger et al. (2014), in which they developed an on-the-job search model to describe worker's career dynamics in terms of income and employment.

By combining these two approaches, we are able to model the dynamics under different wage contracts in a unified framework. The difference between wage contracts is modeled on worker's effort and its effect on production, since performance pay jobs tend to be more skill-driven, and the outcomes are more dependent on the efforts from the employees. We endogenize the choice of wage contracts to account for the intrinsic differences between the two groups of jobs. The key factor in the endogenous choice of wage contracts is the monitoring cost of effort: it captures the difficulties of monitoring and enforcing worker's effort, as well as the measurability of output in order for it to become contractable.

We model the production as a composition of different elements and highlight a key pair: the productivity of routine tasks and the productivity of individual efforts. Routine task is proxied by hours of work, and individual effort depends on the worker's unobserved ability and preference for leisure. Since the use of performance pay is considered to induce a more efficient level of individual efforts, we are interested in studying the implications of growth in different productivities on labor market income distributions.

We add these modeling features to the framework in Bagger et al. (2014). Under this framework, the employment and wage dynamics of each worker are results of on-the-job search for outside options and wage bargaining. The model generates trackable solutions on wage and employment, which can be adapted for model estimation using Indirect Inference. As a result, the combined approach allows us to model the dynamic process under different wage contracts, and study the implications of different production factors on the differences between wage contract groups.

1.3.1 Model setup and timeline

1.3.1.1 Period Production

Firm's production in period t is modeled as follows (in log terms):

$$y_{ijt} = p_j + \theta_i + g(t) + \gamma_{jt}e_{it} + \phi_{jt}h_{ijt} + \varepsilon_{ijt}$$

p_j is the firm-specific productivity for firm j, θ_i is the worker-specific productivity. $g(t)$ is the productivity growth due to accumulated working experience and is assumed to be the same for all workers[8]. The term $\gamma_{jt}e_{it}$ is the productivity associated with worker's effort e_{it}, and γ_{jt} is the return to worker's individual effort due to firm j's production technology. Finally, $\phi_{jt}h_{ijt}$ is the productivity through hours of work: ϕ_{jt} is productivity of working hours that fluctuates over time, and h_{ijt} is the hours of work in each period. ε_{ijt} is a period by period random productivity shock.

Workers' utility is modeled on their preference for consumption, leisure, and the cost of effort. We assume there is no technology for saving, and the workers consume all of their labor income in each period. Hours of work and exerting effort cause disutility for workers. Formally, worker's period utility is given by:

$$U = C_{it} - h_{ijt}^{\omega} - exp(e_{it} - \alpha_i)$$

h^{ω} is the disutility from working (by the hour), and $exp(e_{it} - \alpha_{it})$ is the disutility from exerting efforts. e_{it} is the level of effort and α_i is the worker's ability. A higher level of α_i means it is easier for the worker to exert (the same level of) effort.

[8]Following the literature, we model the productivity growth as a polynomial function of experience: $g(t) = \psi_1 t + \psi_2 t^2 + \psi_3 t^3$

1.3.2 Choice of Contract: Performance Pay and Fixed Wage

We model the choice of wage contracts following the approach used in Lemieux, Macleod and Parent (2009). It is built on the insight from Lazear (1986) that when a worker's ability is not perfectly observed, there could be a mismatch between the worker and the task, and performance pay can be used to reduce the mismatch and increase productivity. Formally, we model a worker's ability α_i as private information to herself, firms could not observe nor verify it. We assume α_i follows distribution $\alpha_i \sim N(\hat{\alpha}_i, \sigma_i^2)$. The distribution is common knowledge to both the worker and the firm. $\hat{\alpha}_i$ σ_i^2 are based on worker's observable characteristics, such as education level, age, and marital status. In a fixed wage contract, workers agree to supply a fixed level of effort $\bar{e_{it}}$; in a performance pay contract, workers would choose the optimal level of effort based on their true ability $e^*_{it}(\alpha_i)$. In order to implement performance pay, there is a monitoring cost per period for the firm, denoted as M_j. It is assumed to be firm-specific and time invariant. The monitoring cost for performance pay is based on the idea that contractibility of output and monitoring cost of effort affect the choice of contract form in employment relationships. In many situations, the output is hard to measure objectively, and hence it is difficult to contract on them - it's hard to compensate workers based on their work when both parties could not agree on the outcome. When a worker's effort needs to be constantly monitored to prevent shirking at work, and when the monitoring cost is high, the benefit of added productivity may be outweighed by the monitoring cost itself. We assume that firms choose the type of wage contracts and make a take-it-or-leave-it offer to worker i. The wage contract is chosen based on known information to the firm so as to maximize firm's expected profit.

Formally, we model the optimal efforts and choice of labor contracts as follows. Each period, the firm chooses the optimal expected aggregate output (production outcome minus worker's period disutility from working hours h_{it} and level of effort e_{it}, and monitoring cost M_j, since a worker's utility in labor compensation c_{it} is linear) because employment contracts are determined through Nash Bargaining, which seek to maximize the total utility of both parties.

Under fixed wage jobs, contracting on effort will be based on the expectation of the worker's ability. The expected aggregate utility related to effort is given by:

$$AU_{ijt}(e_{it})_{FW} = \gamma_{jt}e_{it} - E\{exp(e_{it} - \alpha_i)\} = \gamma_{jt}e_{it} - exp\{e_{it} - \hat{\alpha}_i + \sigma_i^2\}$$

The optimal choice of effort level $\bar{e_{it}^*}$ stipulated in a fixed wage contract is given by $\bar{e_{it}^*}$, and is invariant to the worker's actual ability α_i:

$$\bar{e_{it}^*} = argmax_{e_{it}} AU_{ijt}(e_{it})_{FW} = \log \gamma_{jt} + \hat{\alpha}_i - \sigma_i^2$$

Under performance pay jobs, workers choose optimal level of effort based on their ability α_i. The aggregate output is given by:

$$AU_{ijt}(e_{it})_{PPJ} = \gamma_{jt}e_{it} - exp(e_{it} - \alpha_i) - M_j$$

The optimal level of effort chosen by a worker with ability α_i is given by:

$$e_{it}^*(\alpha_i) = argmax_{e_{it}} AU_{ijt}(e_{it})_{PPJ} = \log \gamma_{jt} + \alpha_i$$

The optimal choice of contract is determined by the aggregate output in the effort-related portion. Under fixed wage contract, the expected output is given by $AU_{ijt}^*(e_{it})_{FW} = \gamma_{jt}[log\gamma_{jt} + \hat{\alpha}_i - \sigma_i^2 - 1]$, under performance pay contract, the ex-ante expected output is given by $AU_{ijt}^*(e_{it})_{PPJ} = \gamma_{jt}[\log \gamma_{jt} + \hat{\alpha}_i - 1] - M_j$. The choice between performance pay and fixed wage contract is then determined by $AU_{ijt}^*(e_{it})_{FW}$ and $AU_{ijt}^*(e_{it})_{PPJ}$. Performance pay contracts will be chosen if and only if:

$$AU^*_{ijt}(e_{it})_{PPJ} > AU^*_{ijt}(e_{it})_{FW} \Rightarrow \gamma_{jt}[\log \gamma_{jt} + \alpha_i] - M_j > \gamma_{jt}[log\gamma_{jt} + \hat{\alpha}_i - \sigma_i^2]$$

$$\Rightarrow \sigma_i^2 > \frac{M_j}{\gamma_{jt}} \tag{1.3.1}$$

For the level of effort $\bar{e^*_{it}}$ in a fixed wage contract, it can be thought of as workers being expected to exert proper efforts at work place and maintain the basic professional standards, such as always be on time for duties or not dozing off at work. These behaviors can be easily observed in workplaces without incurring much costs. For the level of effort $e^*_{it}(\alpha_i)$ under performance pay contracts, either workers' output can be objectively measured (for example, investment bankers making profits on specific deals), or monitored though methods such as internal auditing or annual performance review. In this context, monitoring cost M_j is an abstract term that represents either operational cost for monitoring, or feasibility of measuring performance-based output $\gamma_{jt}e_{it}$.

We use monitoring costs to model the choice of wage contracts for a few reasons. First, previous research shows evidence that the use of performance pay is related to the feasibility of objective evaluation of output or workers' efforts. For example, MacLeod and Parent (1998) shows that when the evaluation of performance are more subjective, firms tend to base compensations on easy measures such as the number of hours worked. Second, even though in reality the choice between performance pay and fixed wage is affected by many other factors, from a modeling perspective such a setup is sufficient to distinguish the key differences between the two contract groups. Workers with higher ability tend to have performance based contracts, such as white collar jobs. It is captured by the distribution of workers' ability. As equation (1.3.1) shows, the higher the noise in the worker's ability α_i is, the more likely a performance pay contract is used. And the optimal effort choices under the two contracts ($e^*_{it}(\alpha_i)$ and $\bar{e^*_{it}}$) show that workers under performance pay are more likely to exert higher levels of effort at work.

1.3.2.1 On-the-job Search Dynamics

We follow Bagger et al. (2014) in modeling the job search process. Workers are randomly matched with firms whose productivities of different dimensions are drawn from an underlying distribution. The wage contract is set through bargaining between the firm and the worker. When a employment contract is signed, all the production uncertainties are unknown, but the distributions of the uncertainties are common knowledge to both parties. When a worker is unemployed and actively looking for jobs, she has probability λ_0 each period of being matched with a firm. When a worker is employed, there is a probability of μ that she permanently leaves the work force, exogenous probability that the employment relationship dissolves (but the worker remains in the work force). Endogenous employment termination happens when there is a random negative productivity shock large enough that makes the value of the contract lower than either party's outside option (unemployment and no production). The probability is denoted as τ. Consistent with the original model in Bagger et al. (2014), we allow for the probability that a newly unemployed worker finds a new job immediately, denoted as κ. When the worker is on the job, at the end of each period, there is probability λ_1 that the worker is matched with an outside firm.

An outside offer from another firm is the main channel through which workers renegotiate wage contracts or switch jobs. Figure 1.1 from Postel-Vinay and Turon (2010) shows the evolution of a wage contract: when the worker receives an outside offer higher than the current employment contract from a less productive firm, she will use this to renegotiate her wage contract and get a larger surplus, since the outside offer becomes her new reservation value. When the outside offer is lower than the current contract value, the worker will hide it from her current employer since it does not help her negotiate up her wage. When the outside firm is more productive than the current employer, the worker will switch to the new job, and use the offer at the previous firm to negotiate a new wage contract.

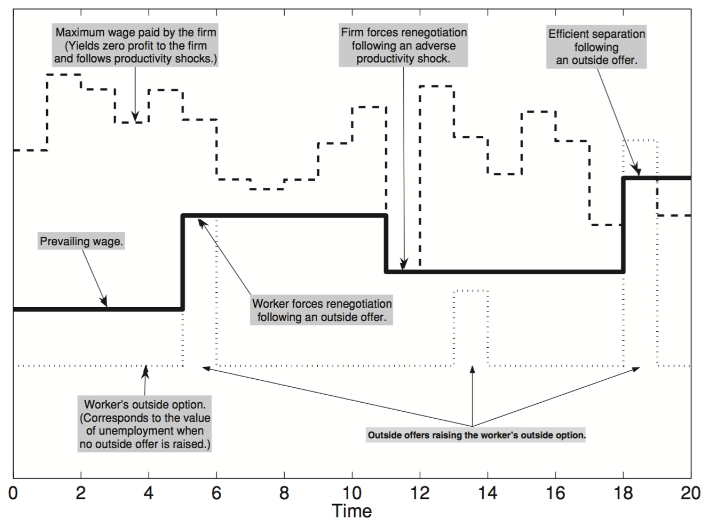


Figure 1.1: Wage negotiation process between firm and workers
**Source:* *Postel-Vinay, Fabien, and Helene Turon (2010) "On-The-Job Search, Productivity Shocks, And The Individual Earnings Process". International Economic Review*

1.3.3 Wage contract and equilibrium model

We use the framework in Bagger et al. (2014) to model a worker's employment and wage dynamics with a few modifications. Their model allows for elegant closed form solutions for a dynamic employment contract, and also has steady state equilibrium solutions for the labor market.

1.3.3.1 Productivity Composition

To use the modeling framework in Bagger et al. (2014), we need to categorize the different productivity elements and transform them into the elements used in their original paper. Their modeling approach includes steady and dynamic productivity elements as well as production uncertainties from different sources. It allows for closed-form solutions of wage contracts and intuitive analysis on the workers' wage and employment dynamics.

The production can be broken down into three elements: 1). Deterministic steady productivities that are known to both parties when the contract is signed, such as firm's productivity p_j; 2). Deterministic productivities that grow over time, such as the worker's productivity growth over experience $g(t)$; 3). Zero-mean productivity uncertainties that vary over time.

Firm's productivity p_j and worker's productivity θ_i are in the first category. We also put the expected productivity from effort, given the endogenous choice of wage contract, into the first category. We assume that the returns to effort γ_{jt} are considered to be stable when signing the contract. This is for modeling convenience so that it allows for closed form solutions to the equilibrium model. It is also used to ensure that workers do not change contracts within each job spell, absent an outside offer. To account for the increase in the return to effort due to technological changes, we allow workers to renegotiate contracts based on their most updated level of returns to effort γ_{it} when they receive outside offers. In this way, effort of production will increase over time within each job spell, and the worker will also be able to negotiate up her wage. Depending on the choice of wage contracts, the return to effort is either $AU_{ijt}(e^*_{it}(\alpha_i))_{PPJ} = \gamma_{jt}[\log\gamma_{jt}+\alpha_i]-M_j$ or $AU_{ijt}(\bar{e_{it}})_{FW} = \gamma_{jt}[log\gamma_{jt}+\hat{\alpha}_i-\sigma_i^2]$. The deterministic growth trend in a worker's productivity is $g(t)$ and it belongs to the second category. The productivity associated with working hours can be divided into two parts: the expected value and a zero-mean random component. The optimal choice of working hours in the contract is given by maximizing the total hours output minus worker's disutility of working:

$$h^*_{ijt} = argmax_h \quad \phi_{jt}h_{ijt} - h^{\omega}_{ijt} = (\frac{\phi_{jt}}{\omega})^{1/(\omega-1)}$$

The productivity from working hours is then given by $\phi_{jt}h^*_{ijt} - h^{*\omega}_{ijt} = \phi_{jt}(\frac{\phi_{jt}}{\omega})^{1/(\omega-1)} - (\frac{\phi_{jt}}{\omega})^{\omega/(\omega-1)}$. Let $\Phi(\phi_{jt}) = E_{\phi}[\phi_{jt}(\frac{\phi_{jt}}{\omega})^{1/(\omega-1)} - (\frac{\phi_{jt}}{\omega})^{\omega/(\omega-1)}]$ and $\xi(\phi_{jt}) = \phi_{jt}(\frac{\phi_{jt}}{\omega})^{1/(\omega-1)} - (\frac{\phi_{jt}}{\omega})^{\omega/(\omega-1)} - \Phi(\phi_{jt})$, then the hours productivity can be broken down into $\phi_{jt}h^*_{ijt} = \Phi(\phi_{jt}) + \xi(\phi_{jt})$. $\Phi(\phi_{jt})$ belongs to the first category as a deterministic steady productivity, and $\xi(\phi_{jt})$

belongs to the second category as a zero-mean productivity uncertainty.

To summarize: the deterministic steady productivity includes firm's and worker's productivity p_j and θ_i, productivity from worker's effort AU_{ijt}, and average productivity from hours of work $\Phi(\phi_{jt})$; the deterministic productivity growth comes from worker's human capital accumulation and skill development $g(t)$; the zero-mean productivity uncertainty consists of period-by-period productivity shock ε_{ijt} and the residuals from hours productivity $\xi(\phi_{jt})$

1.3.3.2 Employment Contract

In our model, an employment contract is defined by three terms: (r, h_t, e_t). r is the proportion of net productivity (firm's production minus worker's disutility from working hours and effort) that the worker has claim to. h_t is the number of hours the worker agrees to work at each period, it varies depending on the productivity shock on working hours each period. e_t is the level of effort the worker agrees to supply, it depends on the form of the chosen wage contract: in a fixed wage contract, worker's effort is chosen at a fixed level $\bar{e^*_{it}}$; in a performance pay contract, worker will choose the efficient effort level $e^*_{it}(\alpha_i)$ based on returns to effort γ_{it} and her own ability α_i.

A worker's labor income consists of two parts: the first part is the direct monetary compensation for her disutility[9] from hours of work and efforts; the second part is the proportional claim to the aggregate output net of monitoring cost if any, as is captured by r in the labor contract. Since output is in log terms, the proportional claim is additive. Formally, the second part wage is:

$$W_{ijt} = \begin{cases} r + y^{PPJ}_{ijt} - M_j + exp(e^*_{it} - \alpha_i) + h^{*\omega}_{ijt} & \text{for performance pay} \\ r + y^{FW}_{ijt} + \gamma_{jt} + h^{*\omega}_{ijt} & \text{for fixed wage} \end{cases}$$

The reason for such a two part wage schedule is that, in Nash Bargaining, both parties will divide based on the joint surplus that are maximized to the best of their ability. A

[9]expected disutility in the case of fixed wage contract

worker's outside option is her reservation wage. In that case, she can enjoy total leisure without exerting any effort. Firm's outside option is shutting down the production. As a result, compensating for worker's working disutility before dividing the surplus is a natural outcome. In essence, a worker's disutility is translated into production cost for the firm in the model. In our model, a worker's compensation varies from period to period under a fixed wage contract. This is consistent with the empirical fact, partly because factors such as overtime work affect a worker's total labor income, and our model specifically accounts for that. The key difference between the two wage contracts in our model is consistent with Lemieux, Macleod and Parent (2009), it is the incentive and compensation structure for worker's effort that distinguish the two types of contracts.

Before modeling the wage contract, we first define a few transformed elements of production based on the three productivity categories defined above. The purpose is to simplify the analytical structure and to be consistent with the modeling framework in Bagger et al. (2014).

We define $\tilde{p} \sim F(p, \theta, M, \alpha, \gamma, \phi)$ as the fixed components of production in a worker-firm match. The distribution of $\tilde{p}$ is determined by the productivity components $(p, \theta, M, \alpha, \gamma, \phi)$. The total value of an employment contract to the worker at period t is given by $V(r, g_t, \tilde{p})$, where $g_t = g(t)$ and $\tilde{p}$ jointly determine the total surplus of an employment relationship going forward. Since r is the worker's proportional claim to total surplus[10], $V(0, g_t, \tilde{p})$ represents the total value of an employment match. When a worker is approached by an outside firm with production characteristic $\tilde{p}'$, the outcome will depend on how it compares to the worker's current employer $(\tilde{p})$. There are three scenarios:

Case 1: $\tilde{p}' > \tilde{p}$

In this case, the worker is approached by a firm with a higher level productivity. The outside firm will be able to poach the worker because they can afford to offer a higher wage than the

[10] e^r is the nominal proportion, since production and wage is in log terms in the model

current firm does. The current employment match value becomes the worker's new outside option, and the wage negotiation with the outside firm will be based on it, because the worker can use the outside firm offer to extract all the surplus from her current employer. Formally, the wage bargaining with the outside firm is as follows:

$$E_t V(r', g_{t+1}, \tilde{p}') = E_t\{V(0, g_{t+1}, \tilde{p}) + \beta[V(0, g_{t+1}, \tilde{p}') - V(0, g_{t+1}, \tilde{p})]\}$$

$\beta \in [0, 1]$ is the bargaining power of the worker, $V(0, g_{t+1}, \tilde{p})$ is the total surplus from current employment, as the worker's outside option. r' is the new wage rate with the outside firm through bargaining. The expectation term E_t is taken over zero-mean production uncertainties $[\varepsilon_{ijt}, \xi(\phi_{jt})]$, as defined in the previous section.

When the outside firm's productivity is lower than that of the current employer ($\tilde{p}' < \tilde{p}$), there are two possibilities. When the total surplus of matching with the outside firm is reasonably high, the worker will be able to use this outside option to negotiate up her wage with the current employer. On the other hand, when the total surplus is sufficiently low, she will stay on the same job and will not be able to renegotiate for a higher wage. Let the cutoff point be $\tilde{q}$:

Case 2: $\tilde{q} < \tilde{p}' < \tilde{p}$

The worker will be able to renegotiate a more favorable wage contract with the current employer. The wage renegotiation will raise the wage to r':

$$E_t V(r', g_{t+1}, \tilde{p}) = E_t\{V(0, g_{t+1}, \tilde{p}') + \beta[V(0, g_{t+1}, \tilde{p}) - V(0, g_{t+1}, \tilde{p}')]\}$$

Case 3: $\tilde{p}' < \tilde{q} < \tilde{p}$

The worker's wage will remain the same.

The cutoff value $\tilde{q}$ is when the wage contract from renegotiation is exactly the same as the old contract:

$$E_tV(r, g_{t+1}, \tilde{p}) = E_t\{V(0, g_{t+1}, \tilde{q}) + \beta[V(0, g_{t+1}, \tilde{p}) - V(0, g_{t+1}, \tilde{q})]\}$$

The worker's reservation value from unemployment benefit is $V_0(g_t)$. The wage when the worker is first employed is given by $E_tV(r_0, g_{t+1}, \tilde{p}) = V_0(g_t) + \beta E_t\{V(0, g_{t+1}, \tilde{p}) - V_0(g_t)\}$.

The above discussion lays out a worker's employment transition process from one job to another. For unemployment, there is the possibility of endogenous and exogenous separation. For exogenous separation, each period there is probability of μ that the worker leaves the employment pool permanently, and probability δ that an employment relationship dissolves. These probabilities correspond to random shocks to employment, such as an accident to the worker or the firm's sudden bankruptcy. For endogenous separation, when a period shock $[\varepsilon_{ijt}, \xi(\phi_{jt})]$ makes the current employment value lower than the worker's reservation wage, the employment relationship will be terminated. Formally, the condition for endogenous separation is:

$$V(g_t) > p_j + \theta_i + g(t) + AU^*_{ijt} + \phi_{jt}h^*_{ijt} - h^{*\omega}_{ijt} + \varepsilon_{ijt} + \beta E_{t+1}V(r, g_{t+1}, \tilde{p})$$

We add the possibility of endogenous separation to the original modeling framework because it adds an additional dimension of differences between performance pay contracts and fixed wage contracts: as performance pay contracts tend to be more efficient, the corresponding employment relationship is relatively more stable since the added productivity could help to offset random negative shocks.

The model allows closed form solutions for the wage process and therefore is very useful when it comes to simulation for workers' income and employment dynamics. For details on the model solution derivation, we refer the readers to Bagger et al. (2014). In the following, we lay out some key elements of the model solutions.

The wage that a worker receives (compensation for disutility and proportional claim to the output) in each period is as follows:

$$\omega_{ijt} = -\int_{\tilde{q}_{it}}^{\tilde{p}_{it}} (1-\beta)\frac{\rho+\delta+\mu+\lambda_1\bar{F}(x)}{\rho+\delta+\mu+\lambda_1\beta\bar{F}(x)}dx + p_j + \theta_i + g(t) + \gamma_{jt}e^*_{it} + \phi_{jt}h^*_{ijt} + \varepsilon_{ijt} \quad (1.3.2)$$

where $F(\tilde{p})$ is the transformed distribution of $\tilde{p} \sim (p, \theta, M, \alpha, \gamma, \phi)$ and $\bar{F}(x) = 1 - F(x)$. One of the key factors in solving the wage equation is the cutoff point $\tilde{q}$ for wage renegotiation. It depends on the current wage rate r and the employment match productivity $\tilde{p}, g_t$. Given the current productivity and cutoff value $(\tilde{p}_{ijt}, \tilde{q}_{ijt})$, the cross-period evolvement of the pair for each worker depends on the probability of an outside offer and random employment shocks. Specifically, the distribution of $(\tilde{p}_{ijt+1}, \tilde{q}_{ijt+1})$ conditional on $(\tilde{p}_{ijt}, \tilde{q}_{ijt})$ is given by:

$$(\tilde{p}_{ijt+1}, \tilde{q}_{ijt+1}) = \begin{cases} (\tilde{p}_{ijt}, \tilde{q}_{ijt}) \quad \text{with probability} \quad 1-\tau-\mu-\delta-\lambda_1\bar{F}(\tilde{q}_{it}) \\ (\tilde{p}_{ijt}, \tilde{q}) \quad \text{with density} \quad \lambda_1 f(\tilde{q}) \\ (\tilde{p}, \tilde{p}_{ijt}) \quad \text{with density} \quad \lambda_1 f(\tilde{p}) \\ (\tilde{p}, \tilde{p}_{min}) \quad \text{with density} \quad (\delta+\tau)\kappa f(\tilde{p}) \\ (0,0) \quad \text{with probability} \quad \mu+\tau+(\delta+\tau)(1-\kappa) \end{cases}$$

where τ is the probability of endogenous separation. The first case refers to the situation where the worker remains on the same job under the same contract. This happens when the worker does not experience unemployment shock, and even when receiving an outside option with probability λ_1, the offer is not good enough to allow renegotiation of wage. The second case refers to when the worker receives an outside offer that is high enough to negotiate a new wage contract but not high enough to switch job. The third case refers to when the worker gets an outside offer high enough that she switches job. The fourth case refers to when the worker becomes unemployed and immediately gets a new job offer. The last case refers to when the worker becomes unemployment in the next period.

1.3.3.3 Labor Market Equilibrium

Following the literature on on-the-job search labor models, we define the labor market equilibrium as a steady state in which the unemployment rate and the density of employment match at each productivity-labor contract pair $(\tilde{p}_{ijt}, \tilde{q}_{ijt})$ to be stable. This implies that the in-and-out flow of the unemployment pool, and the density of workers at each productivity pair $(\tilde{p}_{ijt}, \tilde{q}_{ijt})$ should be the same.

For unemployment rate u, let the proportion of employed workers with experience level t be $a_1(t)$, then the unemployment steady pool is given by the following two conditions:

$$(\lambda_0 + \mu)u(1 - a_1(t)) = (\delta + \tau)(1 - \kappa)(1 - u)a_1(t)$$

$$(1 - u)a_1(t) = [1 - \mu - \tau - (\delta + \tau)(1 - \kappa)](1 - u)a_1(t - 1) + \lambda_0 u(1 - a_1(t - 1))$$

The first equation states that, at each period, the number of previously unemployed workers getting a job should be equal to the number of unemployed workers who lost their jobs. The second equation states that the pool of employed workers with experience t consists of workers from previous period $t - 1$ who remain employed into the next period, and those who had just become newly hired. Together, these two conditions pin down the equilibrium condition for any given level of worker experience.

For the equilibrium conditions of employment at any given level of productivity $\tilde{p}$, consider the distribution of employment density $L(\tilde{p})$. The exit rate of workers from employment productivity less than $\tilde{p}$ is $\mu+(\delta+\tau)(1-\kappa)+((\delta+\tau)\kappa+\lambda_1)\bar{F}(\tilde{p})$. The first half is due to unemployment, the second half is due to job switch when the worker transits to a more profitable firm. The inflow of workers into the pool is given by $\lambda_0 uF(\tilde{p}) + (\delta+\tau)\kappa[1 - L(p)](1-u)F(\tilde{p})$. The first half represents workers who become newly employed, the second half represents workers who are reallocated into the pool because they are poached by outside offers. The equilibrium condition is then given by:

$$\mu + (\delta + \tau)(1 - \kappa) + ((\delta + \tau)\kappa + \lambda_1)\bar{F}(\tilde{p}) = \lambda_0 u F(\tilde{p}) + (\delta + \tau)\kappa[1 - L(p)](1 - u)F(\tilde{p})$$

For a thorough discussion of the solutions to the steady state equilibrium, we refer the readers to Bagger et al. (2014).

1.3.4 Discussions on the model

In this section, we discuss the key features of the model, and the mechanisms through which it captures the labor market income and employment dynamics.

1.3.4.1 Wage Contracts

Wage contract is modeled on uncertainties about worker's ability. As discussed in the previous section, this follows the insight from Lazear (1986) and the modeling approach in Lemieux, Macleod and Parent (2009). The trade off in the choice of contracts is between an inefficient level of effort under fixed wage contracts and an added monitoring cost to enforce effort under performance pay contracts. The choice of wage contracts comes down to three elements: the level of uncertainty on the worker's ability σ_i^2, the productivity of effort γ_{jt}, and the monitoring cost of effort M_j. We assume that the level of uncertainty σ_i^2 is invariant across time, this means that throughout the entire career of a worker, the variation of wage contracts at different jobs depends only on firm level characteristics - productivity of effort and monitoring cost. The use of performance pay contracts depends on the uncertainty: workers with a higher level of σ_i^2 are more likely to choose performance pay contracts. In doing so, we allow for the heterogeneity in wage contracts along multiple dimensions through calibration of parameters at both the worker and firm levels.

1.3.4.2 Job Search and Labor Market Transitions

In our on-the-job search framework, outside job offers arrive randomly in each period. Unlike directed-search models, there is no intrinsic mechanism for the matching and sorting between high productivity firms and workers. Labor market transition for workers is a result of the the search process. Even though the search probabilities are the same for all workers, the model generates differences in transition probabilities through productivities: for job to job transition, workers in a more productive firm are less likely to switch to another job because outside job offers are less likely to beat the current one, since the outside firms' productivities are drawn from the same distribution. For endogenous separation, workers under performance pay contract choose more efficient levels of effort, as a result, they are less likely to terminate the employment relationship under negative productivity shocks.

1.3.4.3 Wage Growth

There are three sources of wage growth in our model. The first one is the increase in worker's productivity through experience $g(t)$ - workers accumulate skills and become more productive on their jobs. The second source is job search - workers use outside offers as leverage to renegotiate higher wage with their current employers, or switch to better-paying jobs. The third source of wage growth is productivity growth over time, including firm-level and worker-level productivity increases, as well as increases in returns to routine tasks (hours of work) and returns to effort. In our model, we assume that these productivities are considered to be static during the contract negotiation process, and productivity growth is modeled as unexpected mean-increasing shock in the underlying distribution.

1.3.5 Model Simulation

Key parameters of the model are simulated based on their corresponding distributions. Firm level productivity p_j follows the Weibull distribution: $F_p(p_j) = 1 - exp(-[\nu_1(p_j - p_{min})]^{\nu_2})$.

Worker's productivity follows (truncated) normal distribution: $\theta_i \sim |N(\bar{\theta}, \sigma_\theta^2)|$. Similarly, γ_j and θ_j follow truncated normal distributions: $\gamma_j \sim |N(\bar{\gamma}, \sigma_\gamma^2)|$ and $\phi_j \sim |N(\bar{\phi}, \sigma_\phi^2)|$. The distributions of γ_j and ϕ_j allow for cross-firm heterogeneities in productivity. To model the within-firm growth in returns to effort and the fluctuation in working hours, we use an additive structure: $\gamma_{jt} = \gamma_{jt-1} + \varepsilon(\gamma)_{jt}$, $\gamma_{j0} = \gamma_j$. The random disturbance $\varepsilon(\gamma)_{jt}$ is strictly positive to make sure γ is increasing over time: $\varepsilon(\gamma)_{jt} \sim |N(0, \sigma^2_{\varepsilon(\gamma)})|$. We model worker's ability α_i based on the average level $\hat{\alpha}_i$ and uncertainty σ_i^2. The cross sectional distribution of $\hat{\alpha}_i$ is drawn from the distribution $N(\bar{\alpha}, \sigma_\alpha^2)$, and the level of uncertainty is linear in worker's average ability: $\sigma_i^2 = \kappa_\alpha \hat{\alpha}_i$. The firm level hour productivity is given by $\phi_{jt} = \phi_j + \bar{\phi}_t + \varepsilon(\phi)_{jt}$. The random shock $\varepsilon(\phi)_{jt}$ follows a zero-mean normal process: $\varepsilon(\phi)_{jt} \sim N(0, \sigma^2_{\varepsilon(\phi)})$, and $\bar{\phi}_t$ is a fixed growth rate for hour productivity. The random shock ε_{it} is assumed to follow an AR(1) process: $\varepsilon_{it} = \eta\varepsilon_{it-1} + u_{it}$ where u_{it} is a mean-zero random shock. Worker's productivity growth $g(t)$ is modeled to be three-order polynomials $g(t) = \psi_1 t + \psi_2 t^2 + \psi_3 t^3$. Monitoring cost is drawn from a uniform distribution: $M_j \sim U(M_{min}, M_{max})$

The key part of transforming the original model in Bagger et al. (2014) into a canonical model that incorporates both wage contracts is transforming the added productivity elements into the three categories as described above. $\tilde{p}_j = p_j + AU^*_{ijt} + \Phi(\phi_{jt})$, the transformed $\tilde{p}$ includes the firm productivity p_j, monitoring cost M_j, returns to effort γ_{jt}. The distribution of $\tilde{p}$ is an empirical distribution from the joint distribution of all the composite elements. In simulation, we construct the distribution of $\tilde{p}$ by simulating its composite elements and add them up. The transformed random productivity shock $\tilde{\varepsilon}_{ijt} = \varepsilon_{ijt} + \xi(\phi_{jt})$ is constructed similarly.

We first simulate worker's employment status based on the job search process $(\delta, \tau, \mu, \lambda_0, \lambda_1, \kappa)$. Then productivity elements $(p_j, \theta_i, g(t), \gamma_{jt}, e_{it}, \phi_{jt}, h_{ijt}, \varepsilon_{ijt})$ are determined through random draws from the underlying distributions of the structural parameters. Worker's income each period is determined by Equation 1.3.2. Through such a simulation process, we are able to generate the wage and employment status for each worker over an extended period of time.

1.4 Structural Estimation

1.4.1 Indirect Inference

We use indirect inference to estimate our structural model. We estimate the model by education groups. There are 3 education groups based on year of schooling: 7-11, 12-14, and 15-20. Each education group has two wage contract subgroups: fixed wage and performance pay.

The theoretical foundations and statistical properties of indirect inference have been extensively discussed in the literature, for example, Gourieroux, Monfort and Renault (1993) and Bruins et al. (2015). Indirect inference has been applied to many fields of empirical research and in particular, structural models. Quite a number of paper on structural labor applied this estimation method, including Bagger et al. (2014) and Altonji, Smith and Vidangos (2013). The basic idea is to select a few reduced form econometric models called "Auxiliary Models" that are of empirical interests, estimate these econometric models from the real dataset, and subsequently find the set of structural parameters such that the model generates a simulated dataset that resembles the real dataset as close as possible along the measures of these auxiliary models. For a more detailed theoretical discussion, we refer readers to Gourieroux, Monfort and Renault (1993) and Altonji, Smith and Vidangos (2013).

1.4.2 Auxiliary Model

In indirect inference, the choice of auxiliary models is as much art as science. In choosing auxiliary models in our estimation exercises, we seek to select aspects of the labor market that are of our main research interests and the focus of our structural model. We use the empirical models presented in Section 2 as auxiliary models for our indirect inference estimations. Namely, labor market mobility in EE and EU transition rates (we use the first differenced estimates on EE and EU for the first three periods, giving us 6 parameters in total), moments on hours of work and wage within job spells (as in Table 1.6. This gives us

8 parameters in total), Mincer Wage Equation estimations (the three coefficient estimates on tenure) and the within-contract group income inequality in distribution and growth rate. In order for the model to reproduce the cross-sectional income distributions and reflect the income inequality at the top percentiles both within and across contract groups, we choose more percentiles at the top: 4 high percentile income levels at 96, 90, 84, 80, and two low percentiles: 60 and 24. We supplement these auxiliary models with additional parameters for the purpose of identification. We include the ratio of performance pay jobs at both the upper and lower half of education levels within each group. We also use the following wage regression on hours and its estimated coefficient on growth trend $t * h_{ijt}$ to capture the increase in hours productivity:

$$w_{ijt} = X_{it}'\gamma + g(t) + \beta_1 h_{ijt} + \beta_2 t * h_{ijt} + \delta_{ij} + \xi_t + \varepsilon_{ijt}$$

where X_{it} represents worker characteristics including age and marital status, $g(t)$ is workers' experience, and t is the linear trend in years[11]. Table 1.6 shows auxiliary model estimates from the PSID data across all groups.

1.4.3 Model Identification

Identification conditions of structural models in indirect inference estimation are difficult to establish as clearly as well-studied methods such as GMM, especially given the complexity and uniqueness of each structural model. The theoretical foundation rests on the mapping between the structural parameters and the auxiliary models. In our case, it means that identification requires one-to-one mapping between the structural parameters of the model and the parameters from the auxiliary models, such that by matching auxiliary models from the empirical data and the simulated data, we can back out the unique set of structural parameters that generate these auxiliary model results. A more formal theoretical perspective

[11] t= year - 1975, where 1975 is the earliest year in the panel.

is discussed in Matzkin (2007), the key identification criteria rests on the "observationally equivalent" condition. Given the difficulties in establishing theoretical identification, Monte Carlo simulation is a practical method to check on the identification of the model locally. For a more thorough discussion of the relevant theories, we refer readers to the indirect inference theatrical literature such as Gourieroux, Monfort and Renault (1993) and Bruins et al. (2015). In addition, Altonji, Smith and Vidangos (2013) has extensive discussion on the use of Monte Carlo simulation to check on local identification conditions.

Although identification of structural models in indirect inference is very difficult to pin down theoretically and not a common exercise in the structural labor literature, there are a few important linkages between our model and the choice of auxiliary models that are crucial in recovering the structural parameters from the PSID data. In the following, we lay out a brief discussion on some of the key aspects of model identification.

1.4.3.1 Identification on Productivity Growth

There are three sources of productivity growth in our model: growth in firm productivity p_j, growth in the productivity of hours ϕ_j and effort γ_j. Since our data spans from 1975 to 1996, we need to disentangle these growth elements. In our data, we do not have firm level information, which limits the modeling of firm level productivity growth. To mitigate this problem, we use the Total Factor Productivity estimates for U.S during that period. See Figure 1.1. For each year of observation in our simulations, we add the corresponding rates of productivity growth as productivity shock. To account for the increase in productivity in hours, we add wage regression on hours to the auxiliary model. As described in the previous section, the coefficient on hour-related wage growth $t * h_{ijt}$ captures the increase in productivity in hours, and by matching the auxiliary model estimates in our simulations, we pin down the hour-related growth rate $\bar{\phi}_t$. Having accounted for the first two elements of productivity growth, we are left with the productivity growth in effort, which is unobservable itself. The Mincer equation estimates in the auxiliary model help pin down the effort productivity growth $\sigma^2_{\varepsilon(\gamma)}$.

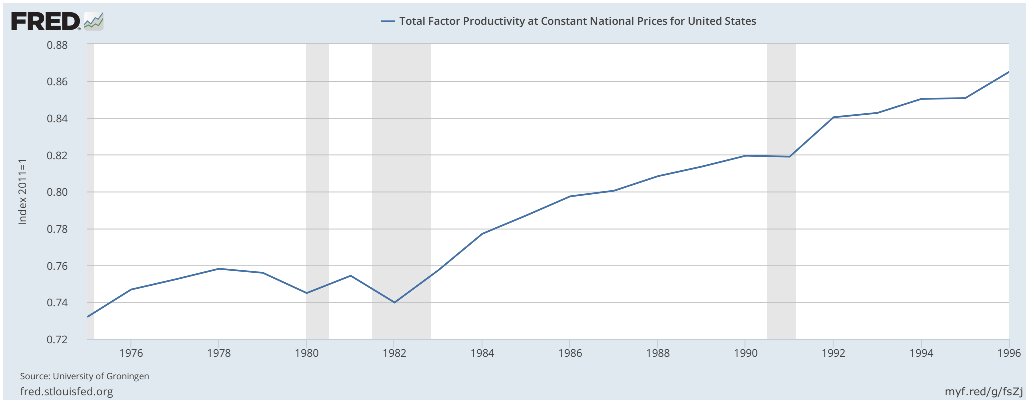


Figure 1.1: Total Factor Productivity. Source: FRED

1.4.3.2 Identification on Wage Contract

The choice of contracts observed in the data helps estimate the distributions of monitoring cost, returns to effort, as well as distribution of worker's ability $(M_j, \gamma_j, \alpha_i)$, as shown in Equation 1.3.1. In our model, monitoring cost and returns to effort $(M_j, \gamma j)$ are drawn from the same underlying distribution for every worker. To allow for heterogeneity on the choice of wage contract at the worker level, we use heterogeneity in uncertainty about worker's ability, σ_i^2. We assume that the level of uncertainty is related to the observable characteristics of the worker: $\sigma_i^2 = \kappa_\alpha * \hat{\alpha}_i$. In our auxiliary model, the ratio of performance pay jobs at different education level helps pin down the uncertainty parameter κ_α. The group differences between fixed wage and performance pay job spells in labor income and employment-to-unemployment rate also helps to pin down the distribution of $(M_j, \gamma_j, \alpha_i)$ because the performance pay contract is inherently better at inducing more efficient level of effort and higher output. It will be reflected in the labor income differences between groups, and also the productivity gain will make employment relationship more robust to negative productivity shocks, which translates into lower separation rate for performance pay jobs.

1.4.3.3 Identification on Parameters and their Distributions

The productivity of hours ϕ_{ijt} is estimated through the auxiliary model on the distribution of working hours. The cross-sectional distribution of hours of work h_{ij} helps us pin down the average level of productivity ϕ_j, and the variation of working hours within job spells helps to estimate the distribution of productivity shock $\varepsilon(\phi)_{jt}$. The Mincer wage equation estimates the average wage growth for workers as they accumulate more working experience, it is related to two factors in the model. First, worker's productivity growth trend $g(t)$ contributes to wage increase over time. Second, workers will gradually negotiate up their wage with the employer through outside offers, as time progresses they are more likely to receive outside offer (each period there is probability λ_1 of receiving an outside offer). The cross-sectional distribution of labor income is helpful for pinning down the average level of productivity elements $(p_j, \theta_i, g(t), \gamma_{jt}, \phi_{jt}, h_{ijt}, \varepsilon_{ijt})$ and their dispersions. The employment transition rates in the auxiliary models, including employment-to-employment, employment-to-unemployment rates are very important in the estimation of transition probability parameters $(\lambda_0, \lambda_1, \mu, \delta, \kappa)$ and the distribution of firm level productivities (p_j, γ_j, M_j). Employment-to-employment transition rates depend on the probability of receiving outside job offers λ_1 and the joint distribution of (p_j, γ_j, M_j), since the worker will only transfer to another job if the new employment match has higher productivity. Employment-to-unemployment transition depends on both the endogenous and exogenous separation rates (τ, δ, μ), and the dispersion of random productivity shocks $\tilde{\varepsilon}_{ijt} = \varepsilon_{ijt} + \xi(\phi_{jt})$ will determine the endogenous separation rate. The unemployment-to-employment transition rates depend on the likelihood of matching rate with outside firm λ_0.

1.4.4 Numerical Implementation

For each routine iteration, we simulate the entire employment and wage conditions for 20000 workers over 50 periods. We use python script to perform all the simulation and optimization routines. Python libraries such as NumPy and SicPy come handy in our estimation tasks, enabling easy coding and fast computing. To speed up the huge amount of computing task in each iteration, we vectorize the simulation process in NumPy, whose data structure enables fast implementation of large scale calculations. To minimize the objective function, we use a combination of optimization algorithms in order to search for the global minimizer and understand how the model behaves over a wide range of parameter spaces. As the structural model creates a nonlinear mapping between parameters of interest and auxiliary models, traditional gradient-based optimization methods such as Newton-Raphson could not be applied. Instead, we reply on non-gradient-based methods such as simplex algorithm Nelder-Mead, stochastic algorithm Basin-hopping, conjugate direction algorithm Powell's method, and stochastic population based method Differential evolution. Methods such as modified Powell and Basin-hopping are useful in searching for proper initial guesses over large parameter space, and Nelder-Mead is ideal for yielding the eventual estimates with local optimization. The SciPy library in Python allows easy implementation of these optimization tasks.

1.4.5 Structural Parameters

The structure parameter set to be estimated is sufficient to completely characterize and simulate each worker's labor market employment and wage in each period. Some of them characterize the underlying labor market structure, such as bargaining power, some of them characterize the uncertainties in the job search process and the probabilistic distributions of productivities. We have the following structural parameters in the data:

Bargaining power of worker, β. Labor market mobility probabilities: (λ_0,λ_1,μ, δ, κ), the endogenous separation probability τ is calculated from the model using given structure parameters. Distribution parameter for firm's productivity: (p_{min}, ν_1, ν_2); individual productivity $(\bar{\theta}, \sigma_\theta)$; individual ability $(\bar{\alpha}, \sigma^2_\alpha, \kappa_\alpha)$; returns to effort $(\bar{\gamma}, \sigma^2_\gamma, \sigma^2_{\varepsilon(\gamma)})$ and returns to working hours $(\bar{\phi}, \sigma^2_\phi, \bar{\phi}_t, \sigma^2_{\varepsilon(\phi)})$. Parameter governing worker's disutility of working ω; support for the (uniform) distribution of monitoring cost: (M_{min}, M_{max}); parameters for the human capital accumulation: (ψ_1, ψ_2, ψ_3). Random shocks to productivity (η, σ_μ^2).

Table 1.7 lists the estimation results for the three education groups. Overall we find the simplex algorithm Nelder-Mead to be the most efficient and better at finding global minimums of objective function. We find the structural parameters to be quite stable across all the education groups. Bargaining power is estimated at about 0.5 for all three groups, this is a bit different from the estimates in Bagger et al. (2014), where the bargaining power is estimated at about 0.3. In addition to using a different labor market dataset, the differences in bargaining power estimates might also be due to the fact that we are not using the empirical firm productivities. Bargaining power is found to be larger in higher education groups, which is consistent with intuition and most empirical findings. Labor market mobility is similar across education groups. Our estimates show that workers with 12 to 14 years of education (in the middle education group) have higher probability of receiving outside offers. This might be due to the fact that higher education groups have relatively more stable jobs while low education group workers often have difficulties finding new jobs. Firm productivity is found to be higher for higher education groups. We also find that the scale of productivity distribution is quite different in low education groups. This probably corresponds to the relatively large income gap at high percentiles for that group. Individual productivity and ability are also found to be on average higher in higher education groups. We also find the cross-sectional dispersion of working hour productivity to be quite low, this is partly due to the fact that the mathematical solution to working hours is quite sensitive to hour productivity, any large disturbance will create big cross-sectional differences in working hours.

1.4.6 Evaluation of Model Fit

Table 1.8 shows the numerical model fit for all the auxiliary model parameters. Overall, our model is able to capture the key aspects of employment and wage dynamics for each education group.

The model is able to generate labor market mobility for the first three initial periods quite similar to the mobility in PSID data, particularly for employment-to-unemployment transition rates. For employment to employment transition rates (job switch), the model has a bit difficulty in generating first period transition rates close enough to the PSID data, especially for fixed wage jobs. One reason is that there often is quite a big gap between the first period and follow-on period transition rates. For example, the 3 period EE transition rates for Education group 7-11 is (0.233, 0.091,0.075), the fitted transition rates from model simulation are (0.186,0.091,0.075). Since our model keeps a constant probability of outside job offers, it is unable to generate drastically different EE transition rates in the first few periods. The slightly decreasing transition rates in our simulated model come from the fact that the employment-worker match gets higher productivity over time, and becomes more stabilized, as a result, it's less likely for workers to find a higher outside option as time moves on.

Our model is also able to generate pretty good fit for the distribution of working hours and income both cross-sectionally and within job match. To simplify numerical estimation procedures, we use the log of annual working hours and income to scale the numbers and make the results less sensitive. The average and standard deviation of job spell income and working hours are pinned down by the mean and standard deviation parameters related to productivity variables: $(\bar{\gamma}, \sigma_\gamma, \bar{\alpha}, \sigma_\alpha, \bar{\theta}, \sigma_\theta)$. The cross-sectional average of within job spell period to period variation of working hours and income are also captured well by the model through the random dispersion structural parameters. However, our model is not able to reproduce the cross-sectional heterogeneity of within job variations in working hours and income, as shown in the table. One reason is that we do not allow for different random dispersion parameters for different individuals, as a result, each job spell has very similar level of within job dispersion in working hours and income. One remedy for this is to allow for heterogeneity in dispersion variables such as $(\sigma_\gamma, \sigma_\theta)$. For example, in Bagger et al.

(2014), the labor market mobility is allowed to be heterogenous across workers by specifying a flexible functional form between individual ability and labor market mobility. To get a better model fit, one could try making desperation variables tied to each individual's ability or productivity. We consider this to be an issue that could be improved upon through future research.

Our model does a pretty good job in capturing the cross-sectional distributions of labor income. We are able to get a close fit for all of the six percentile measures of average job spell annual income. Furthermore, the model is able to differentiate the income level at each percentile for all the contract and education groups: at a give percentile, the simulated income level is higher for higher education groups and performance pay groups, which is consistent with the PSID data. This is a very important aspect of our structural model because the good fit means the model captures the income inequality aspect of the data pretty well, and we can use it to study the effects of various modeling factors through counterfactual analysis. At each percentile. Figure 1.6 plots the simulated distribution of annual labor income, by contract and education group. Our model captures the fixed wage contract group distribution extremely well, while also produces the income gap between performance pay jobs and fixed wage jobs at high percentiles. The model fit is not as ideal for performance pay, since our structural model tends to smooth over income across all the distributions, while empirically there is a pretty sudden jump for performance pay jobs in high income percentiles. Fortunately in our simulated results, the income gap between contract group is indeed increasing at higher percentile.

1.5 Counterfactual Evaluations

With our structural model and estimated parameters, we can evaluate the effects of different factors on worker's labor market experience, and the overall labor market income distributions. To simply numerical results, we look at the education group 12-14. We focus on three elements of the model: 1). Productivity growth; 2). Wage contract; 3). Search intensity. By modifying elements of the structural model, we simulate data under different hypothetical conditions. We are interested in evaluating how these factors affect the labor market income distribution.

1.5.1 Counterfactual: Productivity Growth

First, we evaluate how the change in different productivity factors affects income distributions. We focus on the four productivity elements in the model: firm productivity, worker's individual productivity, productive returns to working hours and effort. We are interested in how the overall growth in productivity affects income distribution, and how the dispersion of productivity translates into income inequality.

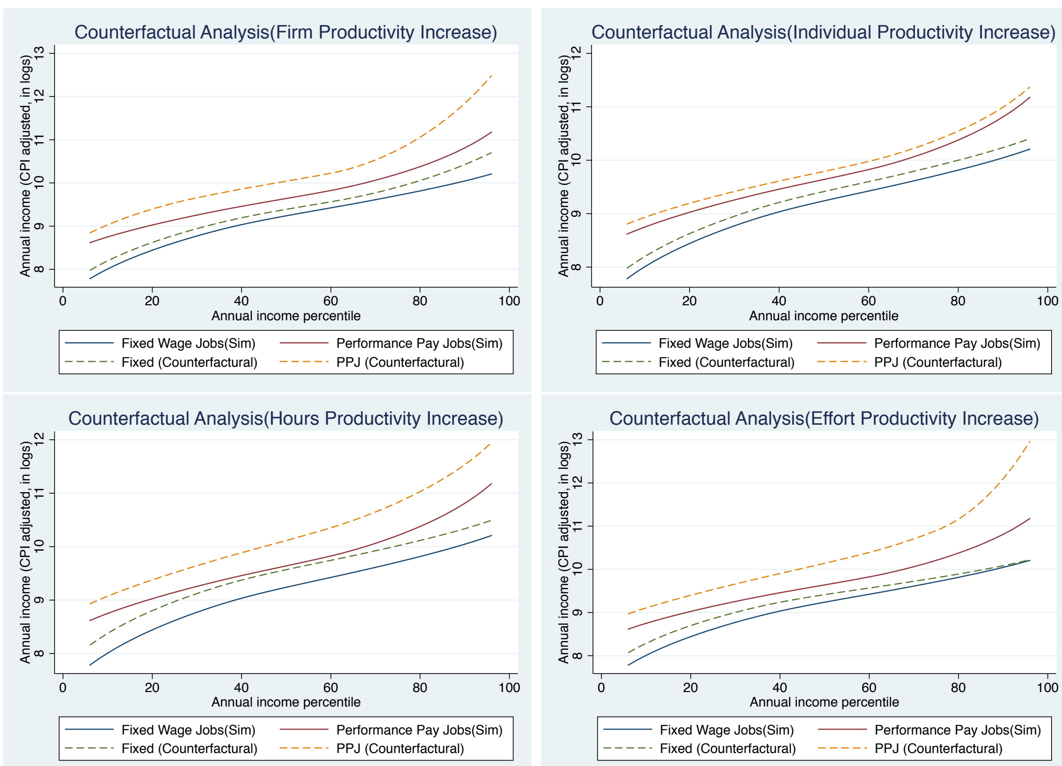


Figure 1.1: Counterfactual: productivity growth and income distribution

Figure 1.1 shows the counterfactual analysis for the growth of productivity. We focus on the effect on income distribution, and in particular, whether different productivity factors have different implications on the income polarization at the top. To simply the analysis, we choose to increase the average level of the productivity distribution by 10%. (correspondingly: $p_{min}, \bar{\theta}, \bar{\phi}, \bar{\gamma}$). The top 2 figures in the graph shows the hypothetical (annual labor) income distribution if there were a 10% increase in average firm productivity or individual productivity, respectively. Comparing these two figures, it's interesting to see the differences in top percentile income: while the increase in individual productivity leads to a uniform increase in income across the entire distribution, the income increases are much larger at the top percentiles for firm productivity increase. This is related to the different roles these factors are playing in our model: individual productivity more or less serves as a fixed level component for each worker, it affects the cross-sectional income distribution only through the production channel. In comparison, firm's productivity also plays an important role in the labor market mobility process: through outside offer, workers will have the opportunity to move up to more productive firms, or negotiate better wage contracts. As a result, an increase in the average firm productivity level is going to allow workers at top income percentiles to gain higher income.

The bottom two figures show the different effects of productivity increase to working hours and effort. The increase in working hour productivity shows a relatively flat increase in income across all the percentiles at the distribution. In contrast, there is a big polarization effect for the increase in effort productivity: there is a much lager increase in income for top percentile performance pay jobs, while for fixed wage jobs the increase is much larger at the lower end. At top percentiles of fixed wage jobs, the increase is almost minimal. This is due to the fact that an increase in returns to effort will lead to a higher likelihood of choosing performance pay jobs, which means that top income fixed wage jobs will be transformed into performance pay jobs, hence a smaller increase at the higher end.

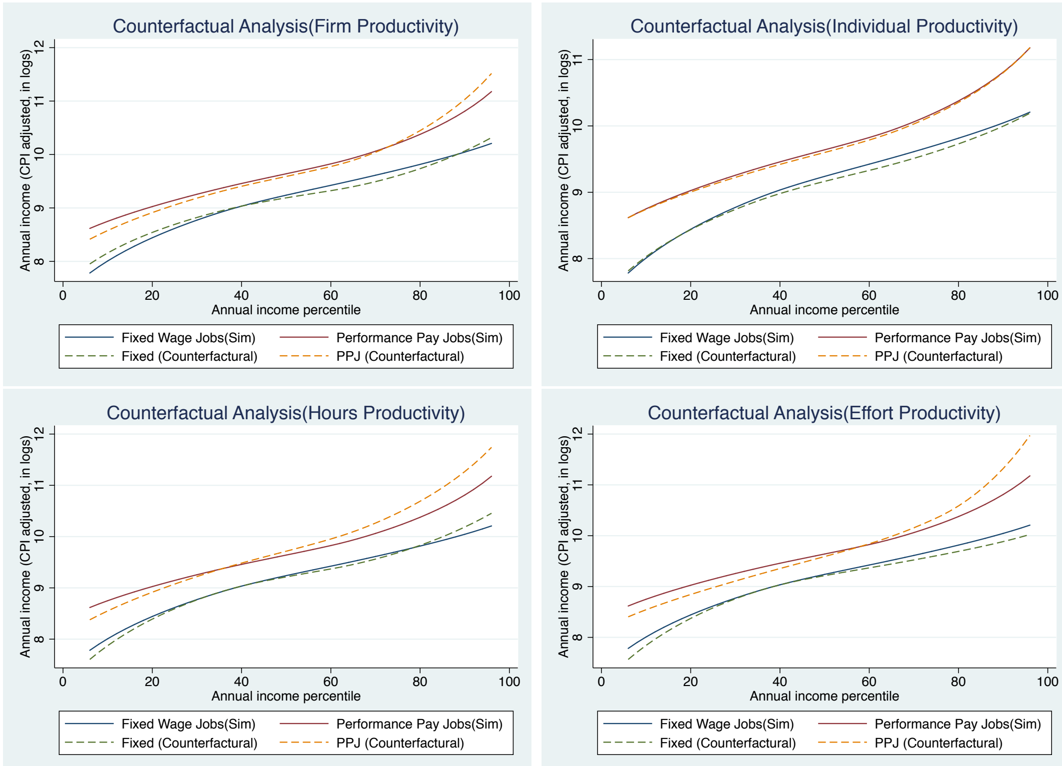


Figure 1.2: Counterfactual: productivity dispersion and income distribution

Similarly, Figure 1.2 shows the counterfactual analysis for increased dispersion of productivity and its effect on income distribution. In our analysis, we increase the standard deviation of these four productivity elements by 10%, based on the originally estimated structural parameters. Similar to the effect of increased productivity level, we see a (slightly) larger income increase at top percentiles for firm productivity than individual productivity. Comparing the returns to hours with the returns to effort, we also find a larger polarization effect in effort productivity.

Our counterfactual analysis on productivity growth shows that changes in different productivity elements have different implications for income inequality, especially considering different wage contracts. Returns to effort have a much larger effect on income polarization at the top end.

1.5.2 Counterfactual: Wage Contract

Second, we evaluate the effect of wage contracts on income distributions. As constructed in our model, the choice of wage contracts depends on monitoring cost, which is a proxy for the feasibility of monitoring worker's effort or measuring worker's output. To that end, we perform counterfactual analysis on two cases: first, we want to evaluate how monitoring cost affects the income distribution in the labor market; second, we want to understand how performance pay as a labor market contracting technology contributes to income inequality. The top two graphs in Figure 1.3 show the results. The left side graph shows the effect of a reduction in monitoring cost on income distribution, and in particular, in comparison to an increase in the productivity of worker effort. We take the baseline distribution of monitoring cost in our estimation for education group 12-14 and reduce the corresponding lower and upper bonds of the (uniform) distribution by 10%.[12] As shown in the graph, monitoring cost affects income distribution of the two contract groups differently. For the performance pay contract group, a reduction in monitoring cost increases income at each percentile for a similar magnitude; for fixed wage contract group, it has little effect on income at low percentiles, while the income at high percentiles has decreased. This is because with the reduction in monitoring cost, the previous top earners in fixed wage group have shifted to using performance pay, and at the same time, low income group has the same income distribution since monitoring cost does not directly affect the payout for fixed wage workers. The comparison with the growth of productivity in effort is even more telling: the productivity growth creates a disproportionate increase in top percentile income for the performance pay group and increases income at lower percentile for the fixed wage group. The almost reverse effects of productivity growth and monitoring cost on income at the two ends of the income spectrum are due to the fact that monitoring cost only serves as a cutoff for the choice of wage contracts, while productivity growth creates a disproportionate increase in the marginal return to effort at the high end. The right side graph at the top half in Figure

[12] from [2.5, 5.8] to [2.2,5.2]

1.3 shows the income distribution when performance pay contracts are shut down and only fixed wage contracts are allowed. Consistent with previous results, the impact on income is larger at higher percentiles of the distribution since performance pay is a channel that translates heterogeneity in effort productivity and worker's ability into income polarization.

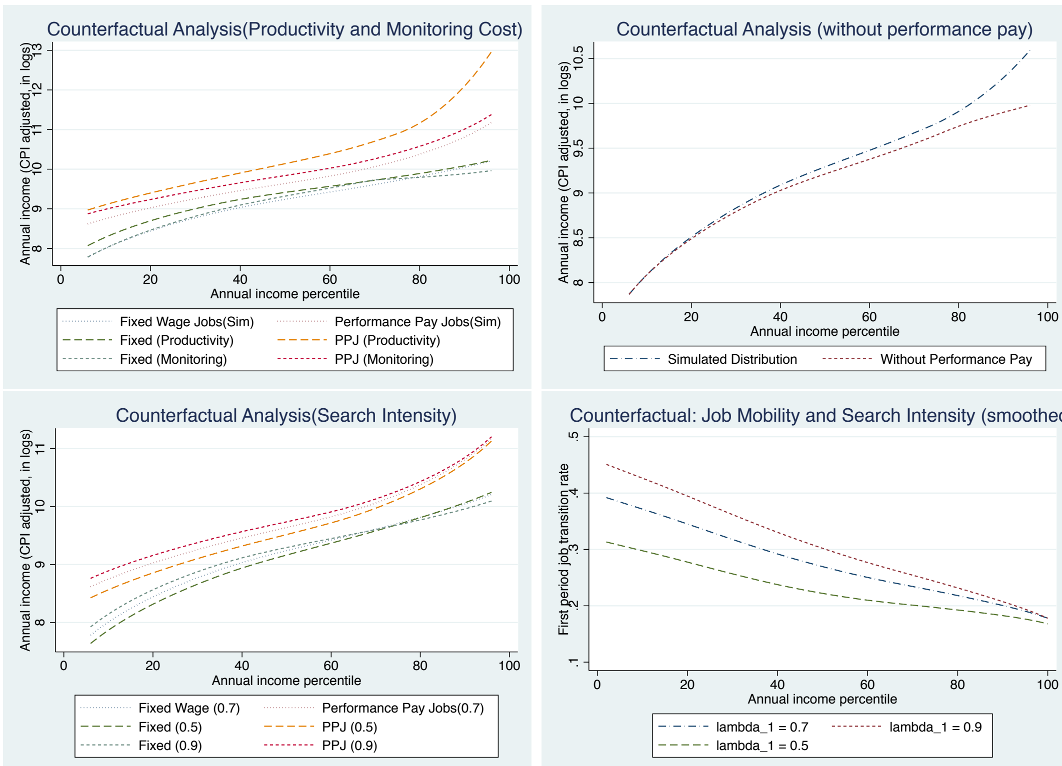


Figure 1.3: Counterfactual: Monitoring Cost, Performance Pay Contract and Search Intensity

1.5.3 Counterfactual: Search Intensity

We also use counterfactual experiments to evaluate the effects of search intensity on income distribution and labor market mobility. In contrast to direct search models, the search intensity in our modeling framework is exogenous. It affects a worker's income mainly through two channels. First, it impacts the worker's wage rate since the possibility of outside job offers will be factored into the wage bargaining process as worker's outside options; second, it affects the worker's labor market mobility through affecting the probability of worker's transition into higher paying jobs. The left side graph at the bottom in Figure 1.3 shows income distribution with different levels of search intensity. We take the estimated intensity $\lambda_1 = 0.7$ as a baseline and experiment higher and lower intensities at $\lambda_1 = 0.9$ and $\lambda_1 = 0.5$. As the graph shows, search intensity has a larger effect on income distribution at the lower percentile. This is because a higher (lower) search intensity has a larger positive (negative) effect on the low income group through both channels: more frequent job search is more likely to yield match with higher paying jobs for low income workers, and the increased matching probability will also be factored into the contract at the time of wage bargaining. The right side of the graph shows the one-period job to job transition rates with different search intensities across the income spectrum. Again, search intensity has a larger effect on job mobility at the lower end of the income spectrum, because random job search implies a higher probability of productivity-increasing matching for low productivity jobs.

1.6 Conclusion

In this paper, we study wage contract and worker's wage and employment dynamics. We find strong evidence in the data that the labor market mobility, returns to tenure, wage growth and working hours are quite different between job spells under fixed wage contracts and performance pay contracts. Furthermore, the labor income gap is concentrated on the top percentiles for performance pay jobs, contributing to both within and across contract group income inequality, especially for low education groups. We further develop a dynamic structural model that incorporates the endogenous choice of wage contracts, and describes workers' labor market experiences over their careers. We then apply the model to structural estimation. Our estimation results suggest cross-group differences in the underlying factors in the labor market, and the model does a good job in explaining important empirical aspects of the labor market data.

From the empirical evidences on the employment and wage dynamics, we highlight the differences across contract groups on labor market mobility: employment relationships are more stable under performance pay jobs, workers are less likely to switch to another firm or being laid off by their employers. The differences are consistent through all education groups and also in the subsample in which workers have experiences under both types of employment contracts.

Our model focuses on wage, working hours, the job change and termination process, as well as cross sectional labor income distribution. We differentiate contracts through flexibility in the wages as well as the uncertainties on productivity: performance pay can link labor income to eventual production outcome, therefore working hours will be more responsive to ex-post productivity shock. Moreover, it creates a channel through which firms and workers can share risks unknown prior to contract signing, making the employment relationship more resilient.

Our structural estimation seeks to uncover group-wise differences underlying the labor market, such as workers' bargaining power, the heterogeneity of firm and worker productivity, and also the magnitude of different productivity shocks, in a bid to offer more insights into the structure of the labor market. Through counterfactual analysis, we are able to show the contribution of different productivity factors to income inequality.

We consider our contribution to the literature in three aspects: to the literature of income inequality, we develop a structural model to evaluate the effects of various labor market factors, and in particular, the role of wage contracts. Our work also brings a new perspective to the research on "skill biased technological change". To the literature on performance pay, we make extensions to the empirical study of labor contracts by exploring the cross contract group differences in workers' employment and earnings dynamics, instead of cross-sectional differences. To the literature on structural labor and on the job search models, we show that it's important to recognize the differences on employment and wage processes between different types of employment contracts, and it's possible to extend existing model frameworks to properly factor in the contractural aspects for structural estimation.

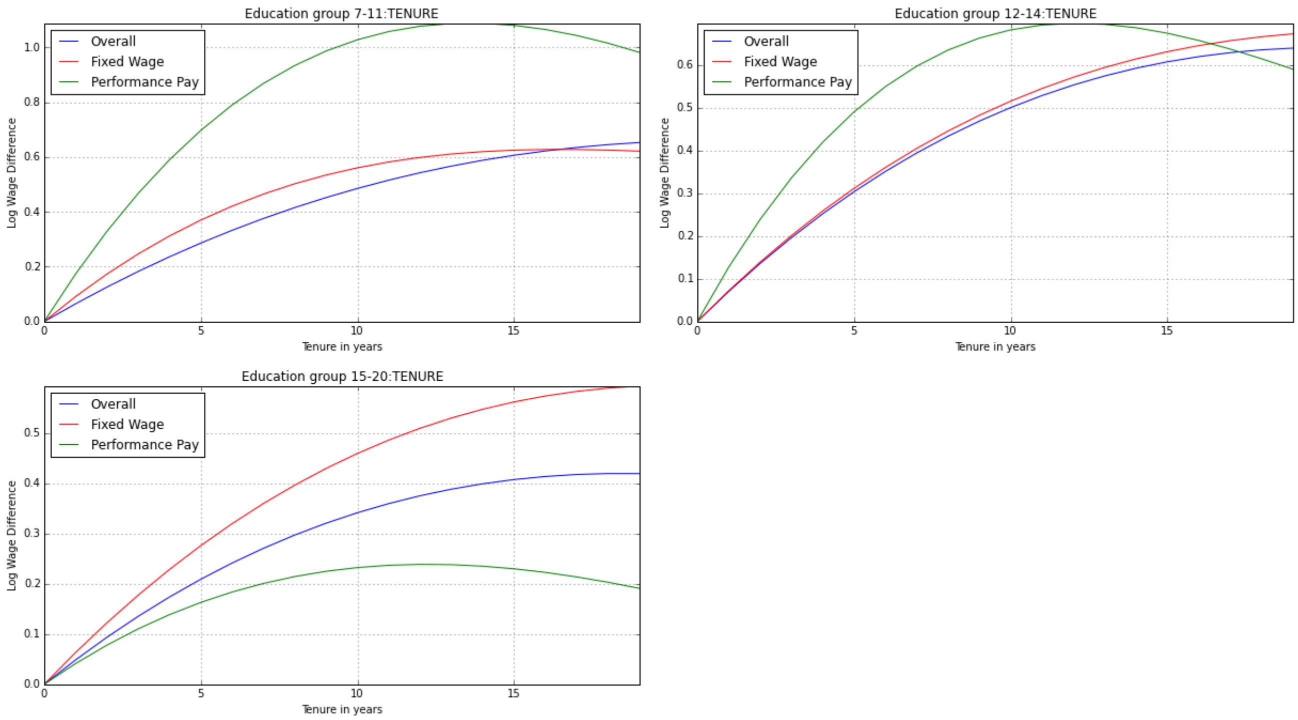


Figure 1.1: Wage-Tenure Profile from Mincer Equations (by education group)

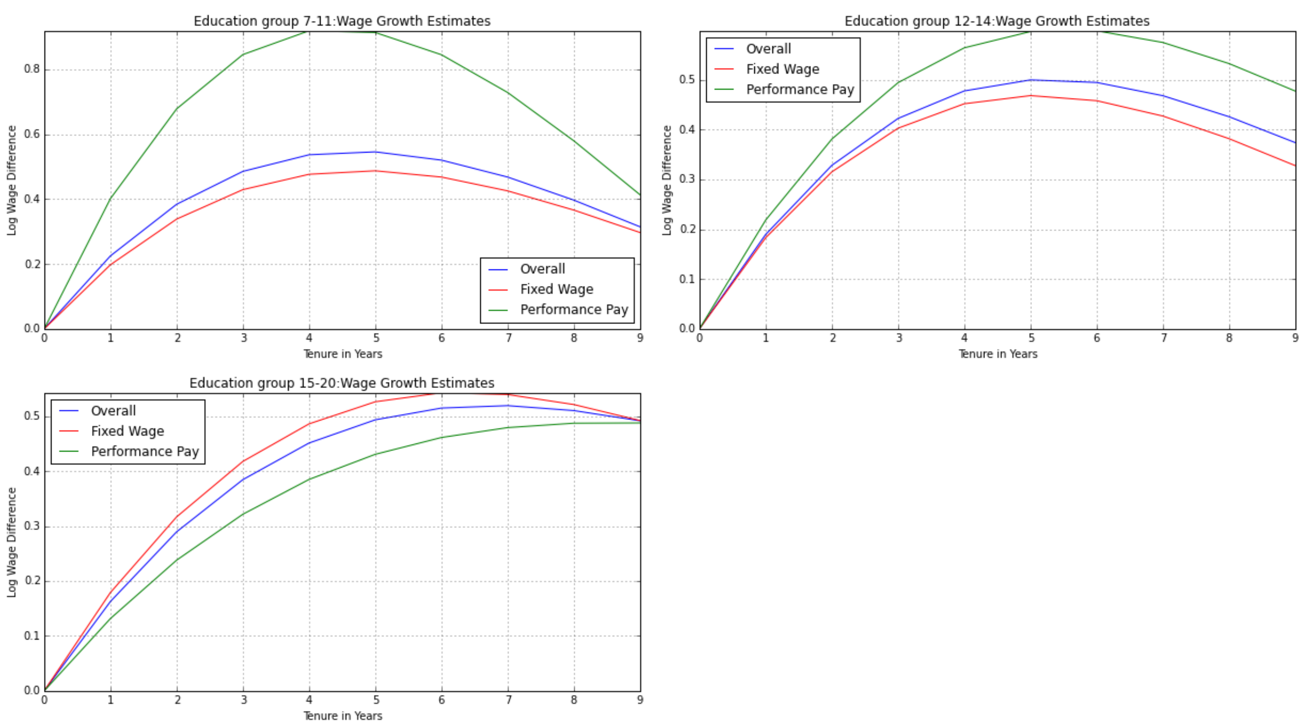


Figure 1.2: Within Job Spell Wage Growth (by education group)

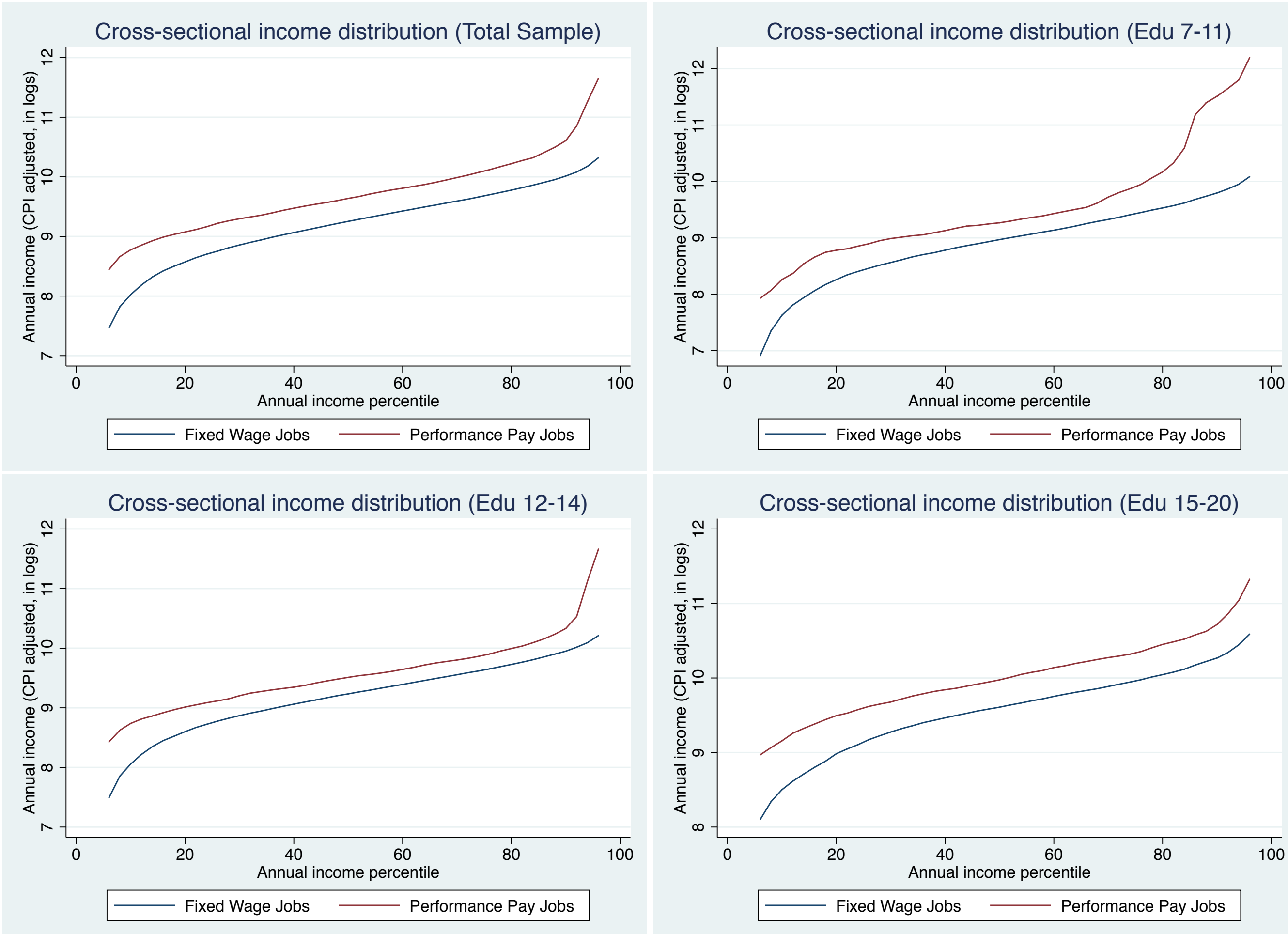


Figure 1.3: Distribution of Income by Contract Group

**Data Source:* *Panel Study of Income Dynamics (PSID).*

The graph shows the income distribution from PSID data on worker's annual income. The four graphs correspond to the entire sample set, and three subgroups by education level, respectively. Each graph shows the distribution of average annual income (in logs) of every job spell, by their wage contract type. To avoid outliers and top coding issue in PSID, we only keep percentiles between 4%-97%.

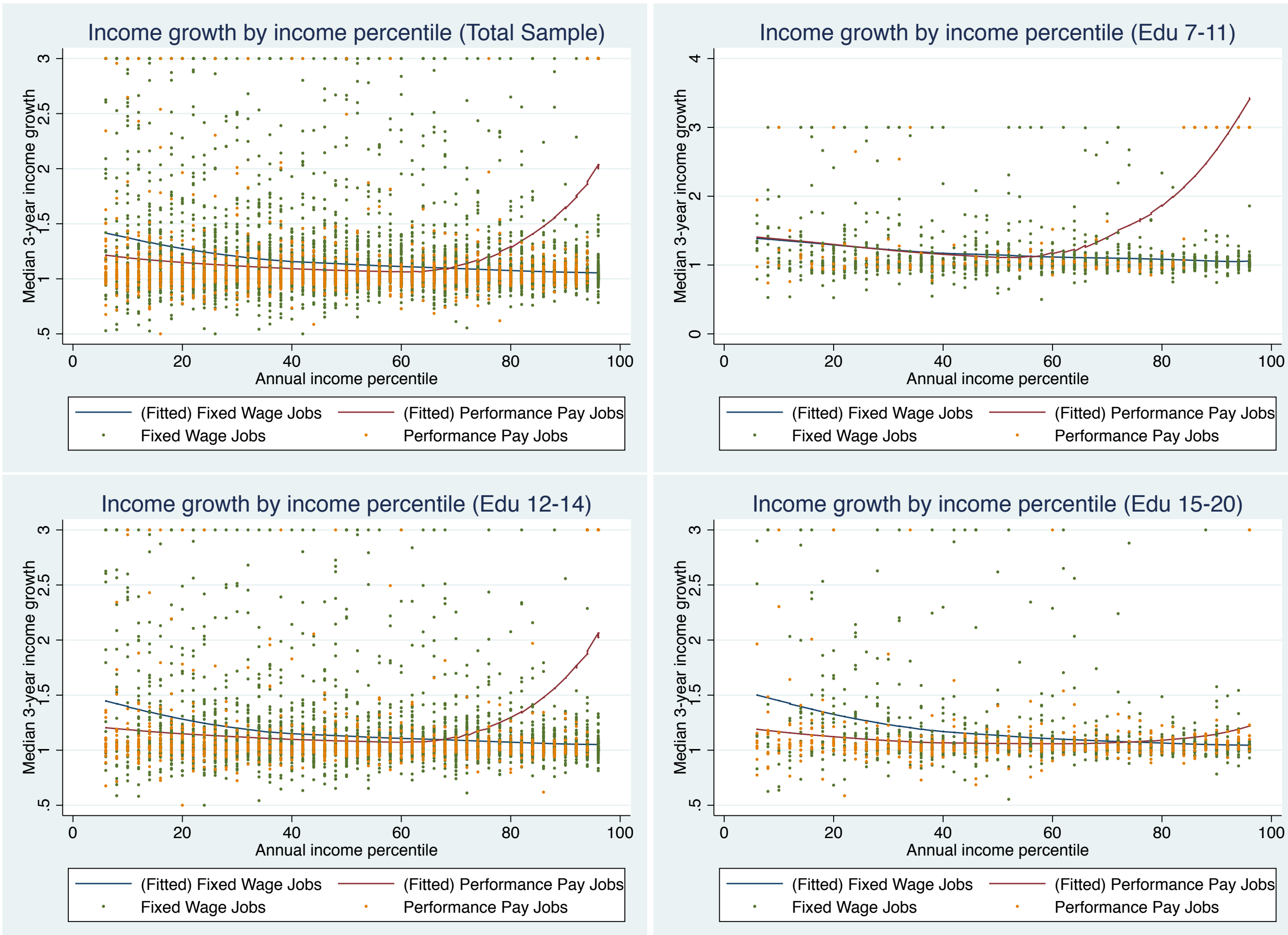


Figure 1.4: Distribution of 2-year Income Growth at Each Income Percentile

**Data Source:* *Panel Study of Income Dynamics (PSID).*

The graph shows the distribution of 2-year labor income growth at each income percentile. The four graphs correspond to the entire sample set, and the three subgroups by education level, respectively. Income percentile is defined by the distribution of the average annual labor income of each job spell in the sample. To avoid outliers and top coding issues in PSID, we only keep percentiles between 4%-97%. Within each job spell, we calculate the income growth rate of any 2-year gap observations, and take the median growth rate for each job spell to account for fluctuations of growth rates within job spell.

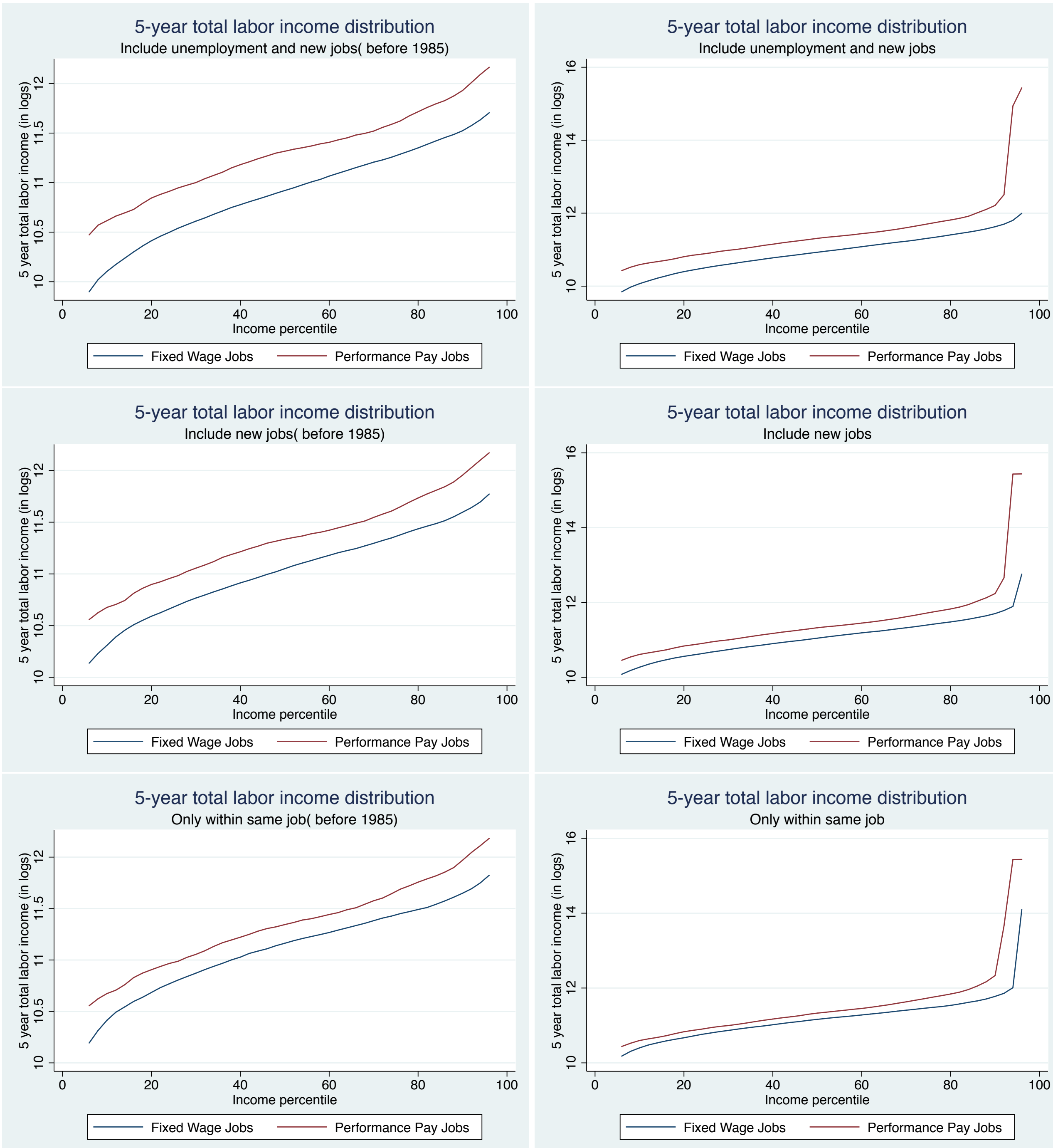


Figure 1.5: Distribution of 5-year Total Income by Contract Group

***Data Source:** *Panel Study of Income Dynamics (PSID).*

The graph shows the distribution of 5-year total annual labor income (in logs) by contract group. To avoid outliers and top coding issues in PSID, we only keep percentiles between 4%-97%. For the three graphs on the left side, we only include job spells that start before 1985. For the three graphs on the right side, we use all the job spells in the sample. For each qualified job spell, we start with the earliest period of observation, and calculate the total annual labor income over the next five years.

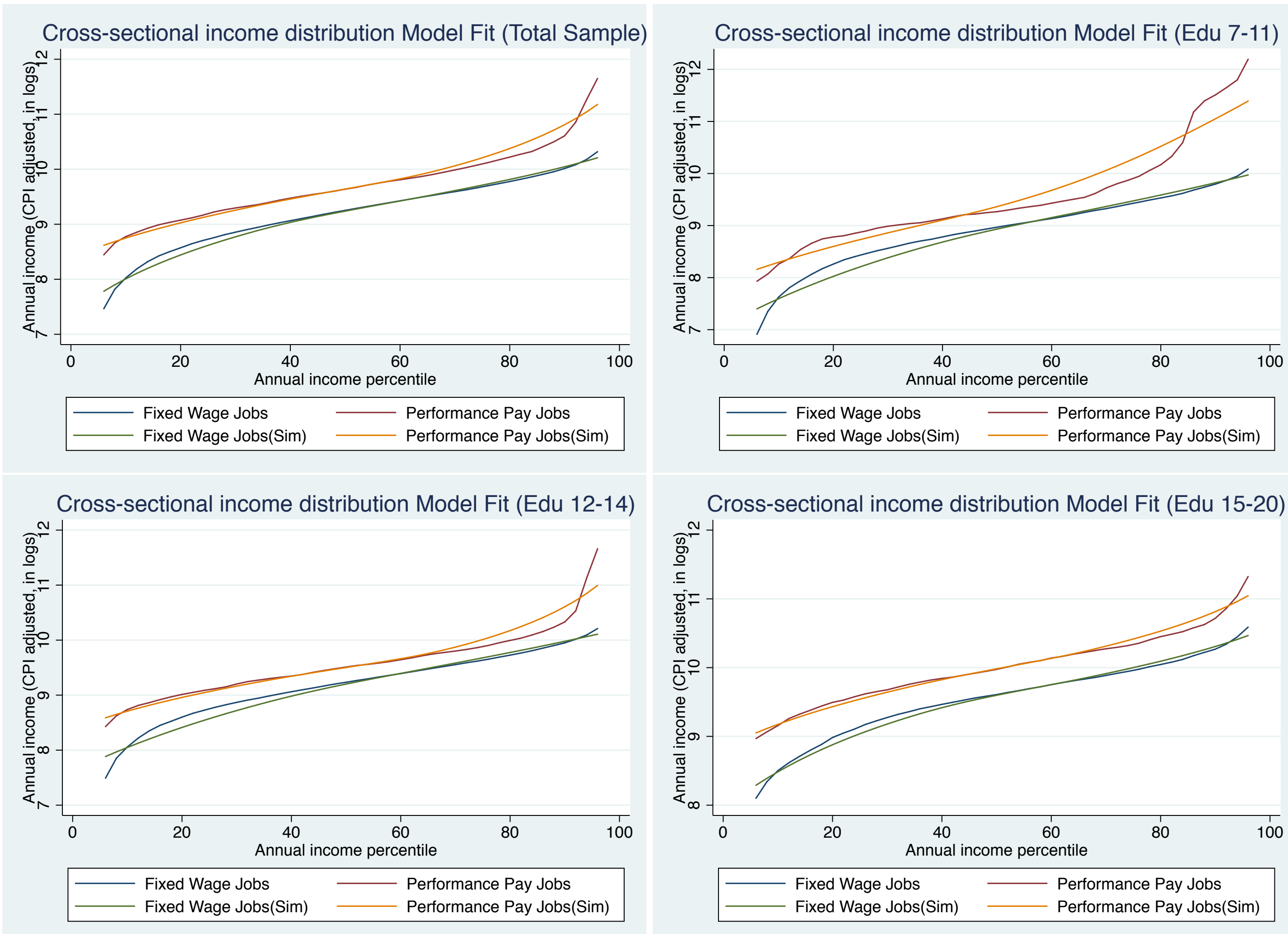


Figure 1.6: Model Fit: Distribution of Income by Contract Group

Data Source: Panel Study of Income Dynamics (PSID).

The graph shows the income distribution from PSID data on worker's annual income and the model fit from simulation. We use estimated structural parameters to generate simulated labor market data and estimate the distribution of annual income.

Parameter	Edu:7-11			Edu:12-14			Edu:15-20		
	Full	Fixed	PPJ	Full	Fixed	PPJ	Full	Fixed	PPJ
$1-\hat{S}_{EE}(1)$	0.2203	0.2332	0.0614	0.2239	0.2501	0.0500	0.2164	0.2594	0.0614
$\hat{S}_{EE}(1)-\hat{S}_{EE}(2)$	0.0866	0.0908	0.0426	0.1031	0.1123	0.0511	0.1044	0.1181	0.0617
$\hat{S}_{EE}(2)-\hat{S}_{EE}(3)$	0.0895	0.0945	0.0407	0.0862	0.0907	0.0638	0.0810	0.0899	0.0559
$\hat{S}_{EE}(3)-\hat{S}_{EE}(4)$	0.0582	0.0611	0.0328	0.0579	0.0604	0.0455	0.0566	0.0580	0.0545
$\hat{S}_{EE}(4)-\hat{S}_{EE}(5)$	0.0386	0.0399	0.0286	0.0483	0.0490	0.0459	0.0418	0.0399	0.0495
$\hat{S}_{EE}(5)-\hat{S}_{EE}(6)$	0.0357	0.0338	0.0556	0.0459	0.0495	0.0268	0.0419	0.0383	0.0549
$\hat{S}_{EE}(6)-\hat{S}_{EE}(7)$	0.0604	0.0623	0.0343	0.0426	0.0425	0.0450	0.0550	0.0500	0.0725
$\hat{S}_{EE}(7)-\hat{S}_{EE}(8)$	0.0311	0.0325	0.0206	0.0275	0.0292	0.0226	0.0230	0.0232	0.0248
$1-\hat{S}_{EU}(1)$	0.1002	0.1066	0.0223	0.0767	0.0874	0.0060	0.0497	0.0625	0.0033
$\hat{S}_{EU}(1)-\hat{S}_{EU}(2)$	0.0535	0.0578	0.0126	0.0466	0.0541	0.0089	0.0365	0.0460	0.0109
$\hat{S}_{EU}(2)-\hat{S}_{EU}(3)$	0.0658	0.0725	0.0073	0.0436	0.0522	0.0051	0.0272	0.0343	0.0104
$\hat{S}_{EU}(3)-\hat{S}_{EU}(4)$	0.0531	0.0574	0.0184	0.0415	0.0487	0.0100	0.0251	0.0300	0.0149
$\hat{S}_{EU}(4)-\hat{S}_{EU}(5)$	0.0504	0.0554	0.0109	0.0343	0.0402	0.0083	0.0296	0.0395	0.0088
$\hat{S}_{EU}(5)-\hat{S}_{EU}(6)$	0.0357	0.0343	0.0488	0.0340	0.0383	0.0160	0.0217	0.0220	0.0281
$\hat{S}_{EU}(6)-\hat{S}_{EU}(7)$	0.1307	0.1407	0.0204	0.0874	0.1002	0.0375	0.0878	0.0986	0.0679
$\hat{S}_{EU}(7)-\hat{S}_{EU}(8)$	0.0217	0.0214	0.0252	0.0121	0.0115	0.0153	0.0189	0.0237	0.0103

Table 1.2: Kaplan-Meier estimates of Labor Market Transition Dynamics

Parameter	Edu:7-11			Edu:12-14			Edu:15-20		
	Full	Fixed	PPJ	Full	Fixed	PPJ	Full	Fixed	PPJ
ξ_{11}	0.0665	0.0955	0.1820	0.0725	0.0744	0.1346	0.0501	0.0652	0.0434
	(0.0068)	(0.0093)	(0.0515)	(0.0031)	(0.0035)	(0.0130)	(0.0042)	(0.0052)	(0.0110)
ξ_{12}	-0.0019	-0.0046	-0.0091	-0.0024	-0.0025	-0.0079	-0.0017	-0.0020	-0.0023
	(0.0003)	(0.0006)	(0.0035)	(0.0001)	(0.0002)	(0.0009)	(0.0002)	(0.0002)	(0.0008)
ξ_{13}	0.0000	0.0000	0.0001	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ξ_{21}	0.0201	0.0121	0.1808	0.0553	0.0567	0.0010	0.1000	0.0911	0.0754
	(0.0162)	(0.0162)	(0.1235)	(0.0058)	(0.0063)	(0.0198)	(0.0083)	(0.0101)	(0.0204)
ξ_{22}	-0.0000	0.0002	-0.0039	-0.0015	-0.0017	0.0014	-0.0023	-0.0022	-0.0013
	(0.0005)	(0.0005)	(0.0042)	(0.0002)	(0.0002)	(0.0007)	(0.0003)	(0.0004)	(0.0008)
ξ_{23}	-0.0000	-0.0000	0.0000	0.0000	0.0000	-0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
N	9789	8632	1157	33718	26856	6862	13336	9202	4134

Table 1.3: Mincer Wage Equation Estimates

Parameter	Edu:7-11			Edu:12-14			Edu:15-20		
	Full	Fixed	PPJ	Full	Fixed	PPJ	Full	Fixed	PPJ
ζ_1	0.2604	0.2282	0.4691	0.2165	0.2101	0.2479	0.1817	0.2010	0.1451
	(0.0185)	(0.0187)	(0.0646)	(0.0078)	(0.0087)	(0.0178)	(0.0109)	(0.0141)	(0.0168)
ζ_2	-0.0367	-0.0317	-0.0700	-0.0281	-0.0282	-0.0308	-0.0197	-0.0227	-0.0139
	(0.0033)	(0.0034)	(0.0108)	(0.0013)	(0.0015)	(0.0027)	(0.0017)	(0.0022)	(0.0025)
ζ_3	0.0012	0.0011	0.0025	0.0009	0.0009	0.0010	0.0006	0.0007	0.0004
	(0.0001)	(0.0001)	(0.0004)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
N	6591	5669	922	23291	17700	5591	9751	6334	3417

Table 1.4: Wage Growth Estimates

Stats	Edu:7-11			Edu:12-14			Edu:15-20		
	Full	Fixed	PPJ	Full	Fixed	PPJ	Full	Fixed	PPJ
mean of Hour Mean	2050.9	2018.6	2292.2	2132.8	2104.4	2246.0	2183.5	2130.3	2302.7
sd of Hour Mean	472.6	464.6	463.2	420.3	411.7	435.1	420.3	418.5	399.8
mean of Hour SD	403.0	386.8	524.1	364.5	355.6	399.9	348.4	355.2	333.0
sd of Hour SD	336.5	325.4	390.5	310.6	305.6	327.6	297.5	295.0	302.7
mean of Log-wage Mean	9.2	9.1	9.5	9.4	9.4	9.6	9.8	9.7	10.0
sd of Log-wage Means	1.0	1.0	1.2	0.8	0.8	0.9	0.7	0.7	0.7
mean of Log-wage SD	0.6	0.5	1.0	0.4	0.4	0.6	0.4	0.4	0.4
sd of Log-wage SD	1.0	0.9	1.3	0.9	0.8	1.0	0.8	0.7	0.9
Number of job matches	1407	1241	166	4636	3706	930	1720	1189	531

Table 1.5: Within job match hour/wage moments

Parameter	Edu:7-11		Edu:12-14		Edu:15-20	
	Fixed	PPJ	Fixed	PPJ	Fixed	PPJ
Labor Market Mobility						
Transition EE1	0.233	0.061	0.250	0.050	0.259	0.061
Transition EE2	0.091	0.043	0.112	0.051	0.118	0.062
Transition EE3	0.095	0.041	0.091	0.064	0.090	0.056
Transition EU1	0.107	0.022	0.087	0.006	0.063	0.003
Transition EU2	0.058	0.013	0.054	0.009	0.046	0.011
Transition EU3	0.073	0.007	0.052	0.005	0.034	0.010
Wage growth						
Mincer tenure1	0.096	0.182	0.074	0.135	0.065	0.043
Mincer tenure2	-0.005	-0.009	-0.003	-0.008	-0.002	-0.002
Mincer tenure3	0.000	0.000	0.000	0.000	0.000	0.000
Working Hours (in logs)						
h mean of mean	7.520	7.639	7.589	7.650	7.605	7.692
h sd of mean	0.384	0.377	0.316	0.325	0.285	0.278
h mean of sd	0.288	0.343	0.228	0.242	0.222	0.190
h sd of sd	0.523	0.587	0.431	0.450	0.398	0.415
Annual Income (in logs)						
w mean of mean	9.152	9.577	9.401	9.613	9.772	10.024
w sd of mean	1.026	1.209	0.818	0.948	0.716	0.783
w mean of sd	0.580	1.075	0.464	0.625	0.426	0.491
w sd of sd	0.978	1.352	0.857	1.059	0.776	0.932
Average Job Spell Income Distribution						
96th percentile	10.085	12.195	10.210	11.661	10.588	11.325
90th percentile	9.795	11.509	9.949	10.330	10.267	10.717
84th percentile	9.618	10.592	9.807	10.093	10.118	10.522
80th percentile	9.530	10.169	9.726	9.994	10.044	10.450
60th percentile	9.135	9.430	9.391	9.644	9.752	10.139
24th percentile	8.405	8.853	8.723	9.085	9.105	9.575
Percentage of ratio of Performance Pay jobs (by education)						
Upper Half	0.075		0.156		0.208	
Lower Half	0.061		0.110		0.205	
Hour Rate Growth						
year*log_hour coefficient	0.091		0.088		0.069	

Table 1.6: Auxiliary Model parameters across groups

Parameter (by Education)			7-11	12-14	15-20
Bargaining Power:	β		0.469	0.516	0.669
Labor Market Mobility:	λ_1	(outside offer)	0.651	0.702	0.582
	λ_0	(getting employed)	0.824	0.838	0.877
	δ	(job separation)	0.065	0.055	0.039
	μ	(leaving labor force)	0.002	0.003	0.002
Firm Productivity:	ν_2	(distribution shape)	1.101	1.760	0.999
	ν_1	(distribution scale)	8.055	3.713	3.648
	p_{min}	(distribution support)	1.242	1.931	2.164
Individual Productivity :	$\bar{\theta}$	(average level)	1.103	1.447	1.390
	σ_θ	(dispersion)	0.866	0.905	0.568
Individual Ability :	$\bar{\alpha}$	(average level)	3.664	4.616	5.168
	σ_α	(dispersion)	1.427	1.658	1.515
	κ_α	(uncertainty)	0.352	0.541	0.397
Returns to Effort :	$\bar{\gamma}$	(average level)	1.464	1.613	1.535
	σ_γ	(dispersion)	0.340	0.574	0.310
	$\sigma_{\varepsilon(\gamma)}$	(growth over time)	0.531	1.024	1.230
Returns to Hours:	$\bar{\phi}$	(average level)	2.964	3.216	3.258
	$\bar{\phi}_t$	(average growth rate)	0.015	0.012	0.016
	σ_ϕ	(dispersion)	0.227	0.221	0.218
	$\sigma_{\varepsilon(\phi)}$	(volatility)	0.325	0.415	0.432
Disutility of Working :	ω	(h_t^ω)	1.427	1.454	1.462
Monitoring Cost :	M_{min}	(lower bound)	3.017	2.554	2.384
	M_{max}	(upper bound)	6.014	5.874	5.841
Productivity Growth $g(t)$:	ψ_1	$(\psi_1 t)$	0.068	0.069	0.095
	ψ_2	$(\psi_2 t^2)$	-0.000	-0.004	-0.006
	ψ_3	$(\psi_3 t^3)$	-0.000	-0.000	-0.000
Random Productivity Shock :	η	$(\varepsilon_{it+1} = \eta\varepsilon_{it} + u_{it})$	0.051	0.138	0.068
	σ_u	$(u_{it} \sim N(0, \sigma_u^2))$	0.564	0.605	0.577

Table 1.7: Structural Parameter Estimation Results, by Education Group

Parameter	Edu:7-11				Edu:12-14				Edu:15-20			
	Fixed	(Sim)	PPJ	(Sim)	Fixed	(Sim)	PPJ	(Sim)	Fixed	(Sim)	PPJ	(Sim)
Labor Market Mobility												
Transition EE1	0.233	(0.186)	0.061	(0.081)	0.250	(0.113)	0.050	(0.094)	0.259	(0.160)	0.061	(0.125)
Transition EE2	0.091	(0.091)	0.043	(0.052)	0.112	(0.111)	0.051	(0.080)	0.118	(0.115)	0.062	(0.083)
Transition EE3	0.095	(0.075)	0.041	(0.320)	0.091	(0.087)	0.064	(0.051)	0.090	(0.077)	0.056	(0.046)
Transition EU1	0.107	(0.100)	0.022	(0.030)	0.087	(0.077)	0.006	(0.007)	0.063	(0.050)	0.003	(0.005)
Transition EU2	0.058	(0.045)	0.013	(0.024)	0.054	(0.042)	0.009	(0.004)	0.046	(0.032)	0.011	(0.021)
Transition EU3	0.073	(0.045)	0.007	(0.013)	0.052	(0.040)	0.005	(0.004)	0.034	(0.031)	0.010	(0.022)
Wage growth												
Mincer tenure1	0.096	(0.056)	0.182	(0.113)	0.074	(0.098)	0.135	(0.075)	0.065	(0.048)	0.043	(0.070)
Mincer tenure2	-0.005	(-0.000)	-0.009	(-0.001)	-0.003	(-0.001)	-0.008	(-0.001)	-0.002	(-0.001)	-0.002	(-0.001)
Mincer tenure3	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
Working Hours (in logs)												
h mean of mean	7.520	(7.843)	7.639	(7.764)	7.589	(7.638)	7.650	(7.636)	7.605	(7.359)	7.692	(7.559)
h sd of mean	0.384	(0.411)	0.377	(0.376)	0.316	(0.291)	0.325	(0.294)	0.285	(0.166)	0.278	(0.188)
h mean of sd	0.288	(0.278)	0.343	(0.297)	0.228	(0.267)	0.242	(0.266)	0.222	(0.246)	0.190	(0.218)
h sd of sd	0.523	(0.093)	0.587	(0.109)	0.431	(0.094)	0.450	(0.096)	0.398	(0.135)	0.415	(0.165)
Annual Income (in logs)												
w mean of mean	9.152	(9.117)	9.577	(9.428)	9.401	(9.431)	9.613	(9.472)	9.772	(9.336)	10.024	(10.045)
w sd of mean	1.026	(1.064)	1.209	(1.044)	0.818	(0.883)	0.948	(0.850)	0.716	(0.561)	0.783	(0.649)
w mean of sd	0.580	(0.295)	1.075	(0.645)	0.464	(0.350)	0.625	(0.466)	0.426	(0.147)	0.491	(0.371)
w sd of sd	0.978	(0.325)	1.352	(0.323)	0.857	(0.303)	1.059	(0.242)	0.776	(0.172)	0.932	(0.284)

Table 1.8: Indirect Inference Model Fit

Parameter	Edu:7-11				Edu:12-14				Edu:15-20			
	Fixed	(Sim)	PPJ	(Sim)	Fixed	(Sim)	PPJ	(Sim)	Fixed	(Sim)	PPJ	(Sim)
Average Job Spell Income Distribution												
96th percentile	10.085	(9.991)	12.195	(11.560)	10.210	(10.107)	11.661	(10.995)	10.588	(10.466)	11.325	(11.045)
90th percentile	9.795	(9.818)	11.509	(10.881)	9.949	(9.978)	10.330	(10.605)	10.267	(10.307)	10.717	(10.813)
84th percentile	9.618	(9.665)	10.592	(10.684)	9.807	(9.854)	10.093	(10.324)	10.118	(10.173)	10.522	(10.633)
80th percentile	9.530	(9.569)	10.169	(10.567)	9.726	(9.775)	9.994	(10.172)	10.044	(10.093)	10.450	(10.531)
60th percentile	9.135	(9.149)	9.430	(9.552)	9.391	(9.395)	9.644	(9.661)	9.752	(9.752)	10.139	(10.134)
24th percentile	8.405	(8.248)	8.853	(8.754)	8.723	(8.541)	9.085	(9.044)	9.105	(9.007)	9.575	(9.520)
Percentage of ratio of Performance Pay jobs (by education)												
Upper Half		0.075		(0.073)		0.156		(0.142)		0.208		(0.213)
Lower Half		0.061		(0.058)		0.110		(0.115)		0.205		(0.196)
Hour Rate Growth												
year*log_hour coefficient		0.09		(0.087)		0.088		(0.089)		0.069		(0.064)

Table 1.9: Indirect Inference Model Fit (cont)

Chapter 2

Overtime Pay Regulations and Their Effects on Workers

Evidence from the 2004 rule change to Fair Labor Standard Acts

2.1 Introduction

The impact of regulations on the labor market has always been a topic of debate. Free market advocates argue that government interventions, such as minimum wage or statutory overtime pay push prices in labor market away from efficient level, increasing the cost of labor and reducing output and employment; workers' rights advocates tend to welcome such regulations that are considered to enhance fair labor compensation. Economic theories have different predictions on the effects of overtime pay regulation. Standard competitive model predicts that hours would decrease for workers since mandatory overtime pay increases the price of labor, and neoclassical labor-demand theory suggests that firms would hire new workers to substitute for overtime labor. An alternative view based on labor contract considers hours of work and compensation as bundled-terms in the labor contract. In response to mandatory overtime pay premium, firms can choose to lower the baseline wage rate so that worker's income and hours of work remain the same. According to this view, overtime pay regulation would have no effect on the labor market. The inconclusive prediction on the effect of the regulation from different theories is often the focal point of policy debate and disagreement.

In this paper, we study the impacts on workers from an amendment in 2004 to the Fair Labor Standards Act (FLSA) about statutory overtime pay premium. Under the new rule, the threshold on wage for exemption from mandatory overtime pay was raised from $250 per week to $455 per week, and the definition for jobs that are eligible for exemption were revised. Using labor market data on workers' income and hours of work that spans the period of rule change, we find evidence that the 2004 FLSA rule change increased the wage and income for workers who gained coverage of statutory overtime pay. We also find an increase in income and overtime pay premium for the group of workers with similar job duties but a higher income that exempted them from statutory overtime pay. Our results suggest that the standard competitive model does not capture well the labor market for overtime work, and government regulations could reduce labor market frictions.

Empirical research on the effects of overtime pay regulation are relatively scarce compared to minimum wage law. Costa (2000) studies the impact of FLSA on weekly working hours during the period 1938-1950. By comparing the hours of work of wholesale trade workers to the hours of retail sales workers [1], she found a 5 percent reduction in standard working week hours. Trejo (1991) evaluates two hypothetical models on the labor market response to rising overtime premium: 1. fixed wage model in which firms keep worker's wage intact and reallocate overtime hours to new employees ; 2. fixed job model in which firms choose to lower the wage rate so that worker would still work the same hours and get the same income. Using data from Current Population Survey (CPS) between 1974 and 1978, he found that wage does adjust lower in response to the overtime pay regulation, but the adjustment is not large enough to offset the regulation. Trejo (1993) analyzes the effect of labor unions on overtime compensation and hours. Using CPS data in 1985, he found that unionization reduces overtime hours but increases the overtime pay premium. In another paper, Trejo (2003) found that FLSA coverage has little effect on overtime hours in industries where FLSA has expanded coverage. Rohwedder and Wenger (2015) studies the issue of violation of overtime rules and found that 19 percent of hourly wage workers were paid less than "time-and-a-half", and 11 percent of salaried works who qualify for overtime pay did not receive any.

The research on overtime pay regulation belongs to the much larger body of research on labor law and labor market institutions. A central question on this subject is how does labor market laws and institutions affect labor market outcomes. Botero et al. (2004) investigates the regulation of labor markets through employment laws, collective bargaining laws, and social security laws in 85 countries. They find that heavier regulation of labor is associated with a larger unofficial economy, lower labor force participation, and higher unemployment. Besley and Burgess (2004) studies the impact of labor regulations on manufacturing growth in India, and find that states which amended labor laws in a pro-worker direction experienced lower output, employment, investment, and productivity. Naidu and Yuchtman (2016) studies 19th century labor market institutions through the perspective of labor market frictions and bargaining over rents between workers and employers, and find suggestive correlations between labor strikes and wages. Naidu, Nyarko and Wang (2016) studies a reform in the United Arab Emirates that relaxed restrictions on employer transitions, and find an increase in incumbent migrants' earnings and firm retention. The existing literature points out the importance of law and institutions to the labor market, and the various mechanisms through which they affect labor market outcomes.

[1] covered and not covered by FLSA, respectively

In this paper, we first discuss a simple adverse selection model for overtime work to shed light on the mechanisms through which mandatory overtime pay might affect workers' hours and income. In our model, workers have heterogenous preferences for leisure over overtime work, which is unknown to the firm. Firms choose to offer a uniform overtime pay which might be rejected by workers, and there is substitution of overtime work across groups of workers with different productivity. Our model gives testable predictions on the income and hours for workers with different overtime pay statute. In our empirical section, we estimate the effects of the 2004 overtime rule change on workers' hours, income, and overtime wage premium. Using data from Survey of Income and Program Participation (SIPP), we construct a panel on workers' income and hours of work that covers the 2004 rule change.[2] We categorize jobs into "Exempt' and "Non-Exempt" groups based on their income level, occupation, and whether they are salary workers or paid by the hour. Based on that, we identify two groups of workers in the data: those who were most likely directly impacted by the new overtime rule (workers whose status change from "exempt" to "non-exempt") and those who had most likely remained exempted from overtime pay throughout the period of rule change. We found that on average there is a 2 precent increase in weekly working hours and a 13% increase in hourly rate for those directly impacted workers. Overtime pay premium on average increased for this group of workers, but the effect is not significant. For workers who were always exempted from overtime pay, we did not find any significant change in working hours, but there is some evidence that overtime pay premium had increased for them.

Our research contributes to the study of overtime pay regulation and more generally labor market laws in two aspects. First, we shed light on the policy debate about the efficacy of overtime regulation on improving worker's welfare by studying the most recent overtime rule change in 2004. The existing literature, to a large extent, focuses either on overtime rules that were enacted several decades ago or effects of the rules without the context of a regulation change, therefore, our research provide a more up-to-date perspective on this issue. As the proposed new overtime rule[3], which

[2] The rule change was officially enacted on August 23, 2004

[3] https:
www.dol.gov/whd/overtime/final2016/

raises the salary threshold for overtime pay exemption to \$913 per week, is halted by the court[4], the discussion on overtime pay regulation is becoming even more relevant. Second, we add to the literature on overtime pay regulation by examining a new dataset and using a new empirical model for studying overtime pay premium. We construct panel dataset on workers' income and hours using different data sources and identify groups that are most likely directly impacted by the rule change. To the best of our knowledge, this is the first panel study on the effects of overtime pay rule. Taking advantage of the panel dataset, we use an empirical model to estimate the "overtime wage premium" and the effect of overtime rule change on the premium. A common problem in labor data for studying overtime pay is that we rarely have direct measures of the wage premium since they are usually lumped together with baseline salary as the total labor income, and in the few cases where such information is available, they are often not accurate enough. We model labor income as the total of baseline salary and overtime compensation, and through variations of hours of work within each job spell and the corresponding variations in income, we are able to identify the overtime pay premium with respect to the baseline hourly wage.

The rest of the paper is organized as follows: in section 2, we lay out the adverse selection model for overtime work, in section 3, we discuss the 2004 FLSA rule change, section 4 shows the empirical models and results, section 5 concludes the paper.

2.2 Model

2.2.1 Adverse Selection and Efficient Labor Supply

We consider a simple adverse selection model for overtime labor supply and compensation. First, we use the model to describe the potential effects of a mandatory overtime pay on the labor supply and wage for workers that are directly impacted by the law change. Then we extend the model to incorporate the overtime labor demand and supply for workers of different income tiers. In particular, workers whose overtime pay statues are not directly affected by the law. In doing so, we are able to explore the mechanisms through which mandatory overtime pay rule change for one particular group of workers could have spill-over effects on other groups of workers.

[4]https:
www.dol.gov/whd/overtime/final2016/litigation.htm

In our adverse selection model, firms expand production by having workers work extra hours. At the same time, workers always prefer leisure to working if they were paid the same regardless. The central issue here is the classic asymmetric information problem: worker's true leisure preference is often unknown to the firm. In order to gain as much surplus as possible, firms will have to decide the uniform overtime compensation to the worker without knowing each worker's true type. Our model predicts a few different scenarios for the overtime work supply schedule. As common in the adverse selection problem, it's possible for labor market inefficiency to rise when workers with higher preference for leisure gets priced out. The role of mandatory overtime pay in our model is to provide a "floor" for overtime compensation, and the model provides predictions on overtime work and wage rate that can be tested. In our model extension to multi-segments of workers, firms can assign overtime task to workers of different productivity and overtime pay statue. In this way, we explore a demand channel through which changes in mandatory overtime pay rule for one particular group of workers could have a spill-over effect on other groups of workers. The model provides testable predications on the overtime wage and labor supply for different groups.

2.2.2 Adverse Selection on Overtime Labor Supply

We first consider an adverse selection model on overtime compensation and labor supply for a homogenous group of workers with identical productivity but different preferences for leisure. We consider a perfectly competitive labor market where workers are always compensated by their reservation value for regular hour work, which is $\bar{\omega}$. This allows us to focus on the labor market for overtime working.

Firms gain overtime output by θ. The productivity (demand) level θ is common knowledge. Workers dislike working overtime with disutility $\gamma \in \{\gamma_L, \gamma_H\}$. Same as firm's productivity, worker's overtime disutility (γ_L, γ_H) and their corresponding probability $(1 - p_\gamma, p_\gamma)$ are common knowledge to both parties. We exclude trivial cases where it's never efficient to work overtime $(\theta < \gamma_L < \gamma_H)$. The difficulty in contracting overtime work lies in information asymmetry on worker's side. Firms face production uncertainty as it is difficult to anticipate demand in advance, oftentimes they only find out if extra working hours from workers is necessary at the very end. On top of that, even if firms know the exact production demand for overtime beforehand, they have an incentive to hide this information from workers so as to lower the wage. For workers in a competitive labor market, the high type workers ($\gamma = \gamma_H$) have an incentive to hide their preference to avoid being edged out by other workers with a lower preference for leisure ($\gamma = \gamma_L$). Therefore, we set up the model in which overtime working is not specified in the labor contract. Instead, firms will make take-it-or-leave-it offers to workers for overtime working whenever there is a demand. Formally, firms would observe the overtime productivity at the end of each period to decided if they need to have workers work overtime. When the overtime productivity is higher than worker's reservation value – disutility from working overtime, firms will request workers to stay on the job and work extra hours. The overtime compensation will be chosen so as to maximize firm's profit.

When $\gamma_L < \theta < \gamma_H$, the overtime productivity is less than disutility of overtime working for high type workers, there is no reason for the firm to expect high type workers to work overtime since any profitable compensation offering will be rejected by the firm, therefore, firm will choose to offer $\omega_{OT} = \gamma_L$ as overtime pay, and only low type worker will choose to work.

When $\gamma_L < \gamma_H < \theta$, it is always efficient to work overtime, since overtime productivity is always higher than disutility of working overtime. Firm have two choices on overtime compensation:

Choice 1. Offer $\omega_{OT} = \gamma_L$ as overtime pay, workers of type γ_H will decline to work overtime, low type workers γ_L will work overtime. Firm's expected overtime profit depends on the probability that a worker is of lower type: $\Omega(\gamma_L) = (1 - p_\gamma)(\theta - \gamma_L)$

Choice 2. Offer $\omega_{OT} = \gamma_H$ as overtime pay, all the workers will agree to work overtime, firm's expected profit from overtime is $\Omega(\gamma_H) = \theta - \gamma_H$

Firms will choose to offer overtime wage that generates the highest profit. If $\Omega(\gamma_L) > \Omega(\gamma_H)$, firms will choose overtime wage γ_L, and vice versa:

$$\omega_{OT} = \begin{cases} \gamma_L & \text{if} \quad p_\gamma \theta < \gamma_H - (1 - p_\gamma)\gamma_L \\ \gamma_H & \text{if} \quad p_\gamma \theta > \gamma_H - (1 - p_\gamma)\gamma_L \end{cases}$$

The choice of overtime wage depends on three key factors: 1). the higher proportion of high type workers (p_H), the more likely firm will offer high overtime compensation (γ_H) to encourage overtime working; 2). the higher overtime working productivity (θ) is, the more likely for overtime compensation to be high, since the profitability of added overtime working is more than enough to cover for higher overtime pay; 3). the smaller the gap of overtime working disutility ($\gamma_H - \gamma_L$) is, the more likely firms will offer high overtime wage, because firms only need to give up relatively small level of rent to low type worker (γ_L) to induce high type worker (γ_H) to work overtime.

As a first-best efficient benchmark, workers should always be compensated properly to work overtime whenever the productivity is higher than worker's disutility ($\gamma < \theta$). The adverse selection could lead to inefficiency because the firm is essentially acting as a market monopoly and set uniform price for overtime work. It is therefore possible for high type workers to be priced out of the market even though it's efficient for them to work overtime.

2.2.3 Adverse Selection with FLSA Overtime Pay

The FLSA on overtime pay requires that eligible workers be paid time and a half of regular wage rate for overtime work. In our model, this means that there is a lower bound to overtime pay that firms can offer to workers. For simplicity, we denote the time and half overtime pay as $\omega_{FLSA} = 1.5\bar{\omega}$. The overtime wage offered by the firm would then have a lower bound: $\omega_{OT} > \omega_{FLSA}$.
The effect of such an overtime law on labor market outcome depends on how it affects the supply and demand of overtime working. Firms will demand overtime from worker so long as the productivity is higher than the mandatory overtime pay ω_{FLSA}; workers will supply overtime work if and only if they are properly compensated for the extra hours of working. Here in the analysis, we only consider the non-trivial case where the mandatory overtime pay is higher than low type overtime wage ($\omega_{FLSA} > \gamma_L$). To that end, there are three basic scenarios for the impact of mandatory overtime pay on overtime working.

Case 1: FLSA mandatory overtime pay has no impact on working hours.
When the mandatory overtime wage is lower than overtime productivity, firm will always find it profitable to have workers work extra hours, and if the mandatory overtime wage is not high enough

to induce high type worker to work, the law will not affect worker's overtime supply. Overall, the effect of the law is a welfare transfer from firm to workers.

Case 2: FLSA mandatory overtime pay decreases working hours.
When the mandatory overtime wage is higher than overtime productivity, overtime working becomes inefficient for both parties. Firms will decrease their demand for overtime since they will have to pay at a rate higher than what they could gain from it. In this case, overtime pay regulation will push wage to outside of the efficient price range, and there will be no market for overtime working.

Case 3: FLSA mandatory overtime pay increases working hours.
When the mandatory overtime wage is lower than overtime productivity but higher than high type worker's disutility, it's possible to increase overtime work efficiency if high type workers were priced out of market absent the overtime pay mandate. In this case, overtime pay regulation not only transfers welfare from firms to workers, but also increases overtime working to a more efficient level.

Table 2.5 summarizes all the possible cases. The impact of FLSA overtime pay will depend on how it affects the existing structure of supply and demand for overtime work. It could lead to less efficiency by pushing overtime rate high enough so as to make it unprofitable to the firm; or it could merely function as a welfare transfer to move the rent from firms to workers while maintaining the overtime work structure. Alternatively, it's also possible to improve market efficiency by alleviating the adverse selection problem for monopolistic firms. Our model, parsimonious as it is, is able to capture the information asymmetry problem in the labor market for overtime work, and derive testable alternatives to help understand the effect of the statuary overtime pay premium in FLSA. Based on the model predictions, we will focus on its effect on working hours to see which scenario best explains empirical evidence.

2.2.4 Adverse Selection on Multi-Segment of Workers

Now we consider an extension to the model where we include multi-segments of workers with different productivity and overtime pay statue. There are three types of workers based on their income range: (L, M, H). Workers of type M are directly affected under the overtime pay rule change, group L and H refers to workers whose income are below and above group M, respectively. For firms' demand of overtime work, we consider 4 types of tasks: $\underline{\theta}$ can only be assigned to low group L; $\bar{\theta}$ can only be assigned to high group H; θ_L can be assigned to group L and M, the productivities are θ_{LL} and θ_{LM} respectively; θ_H can be assigned to group M and H, the productivities are θ_{HM} and θ_{HH} respectively. We also assume that worker's income group is tied to her productivity: for the same task, workers in a higher income group can produce more ($\theta_{LL} < \theta_{LM}$ and $\theta_{HM} < \theta_{HH}$). For each group of workers, we consider two sub-groups with different levels of preference for leisure: l and h. Consistent with our baseline model in the previous section, we assume that worker's preference for leisure is unknown to the firm. We denote the preferences for leisure by worker group and type. For example, low type worker of group M delegated to task θ_H has reservation value γ_M^l. We assume that high type worker always has a higher reservation value. ($\gamma_M^l < \gamma_M^h$). The distribution associated with high and low type workers are $(1-p_L, p_L)$, $(1-p_M, p_M)$, $(1-p_H, p_H)$ for the three groups L, M and H, respectively, and they are independent of task. Our model extension aims to explore the substitution of overtime work across different types of workers: overtime task θ_L and θ_H can be delegated to different worker groups, and change in overtime wage for one group of workers has the potential to shift overtime demand from one group to another.

First, we consider task θ_H. It can be delegated to group H or M. Firms can either make an overtime compensation offer to H or M group worker. If they offer it to H group worker first, some workers might reject the overtime offer if the compensation is below their reservation value (γ). In this case, firms will turn to M group workers and make them an offer that maximizes firms' profits. The situation is similar if firms choose to offer to group M workers first. Table 2.1 lists the six scenarios and the corresponding firms' profits.

Depending on the payoff structure, it's possible that firms will delegate high level task θ_H exclusively to group H or M. Or alternatively, firms can choose to prioritize group H or M and make sequential overtime offers to the other group only when the initial overtime offer is rejected. As shown in Table 2.1, when firm's sequential offer is (γ_H^l, γ_M^l) or (γ_M^l, γ_H^l), there's a probability that overtime demand not being met. Such inefficiency arises when workers with high preference for leisure gets priced out due to adverse selection by the firm.

When overtime mandatory pay is enacted for group M workers, firms' overtime offers for low type workers will have to be raised to ω_{FLSA}. We focus on the case where mandatory pay is in-between the high type and low type workers' reservation value ($\gamma_M^l < \omega_{FLSA} < \gamma_M^h$). Two inefficient offers will be directly impacted by the rule change: (γ_H^l, γ_M^l) and (γ_M^l, γ_H^l). In addition, the rule change would lower the corresponding payoffs for firms. Because the change decreases the payoff for a few options of overtime compensation, it would only have an effect on equilibrium overtime hours and compensation if the previous optimal offer by the firm was rendered less profitable than another option. Table 2.2 lists the choices of overtime offer for task θ_H once the overtime mandatory pay was enacted. Here we focus on two interesting cases where the law have different implications for these two groups of workers.

Case 1. when firm's initial optimal offer is (γ_M^l, γ_H^l), as the effect of overtime law kicks in, it's possible for the firm to choose offer (γ_H^l, γ_M^h) [5]. In this case, conditional on a pending task θ_H, group H's overtime probability will increase from $p_M(1-p_H)$ to $(1-p_H)$, and group M's overtime probability will depend on the relative magnitude of $(1-p_M)$ and p_H. As a result, we should expect to see an increase in overtime hours for group H. For such a case, efficiency is improved in the sense that overtime demand will always be met with supply[6], because without mandatory overtime pay, market will choose to offer low wage (γ_M^l and γ_H^l) due to adverse selection, and the high type workers are excluded.

Case 2. When the initial optimal offer is (γ_M^l, γ_H^h), it's possible that firms will offer (γ_H^l, γ_M^l) after the mandatory overtime pay[7]. In this case, conditional on a pending task θ_H, group H's overtime probability will change from p_M to $(1-p_H)$, and group M's overtime probability will decrease from $(1-p_M)$ to $p_H(1-p_M)$. This is a scenario where the introduction of overtime pay does not affect the adverse selection on high type workers in the targeted group (γ_M^h). Instead, firms switch the offer priority from group M to H. In this sense, there is substitution of high type workers (H) for medium type (M) because of the increase in overtime wage. An interesting result is that the change in working hours is ambiguous for high type workers, even though the firm has substituted H as first priority group, they might choose to exclude high type workers (γ_H^h) because they can

[5]when $(1-p_M)(\theta_{HM} - \gamma_M^l) + p_M(1-p_H)(\theta_{HH} - \gamma_H^l) > (1-p_H)(\theta_{HH} - \gamma_H^l) + p_H(\theta_{HM} - \gamma_M^h)$ and $(1-p_M)(\theta_{HM} - \omega_{FLSA}) + p_M(1-p_H)(\theta_{HH} - \gamma_H^l) < (1-p_H)(\theta_{HH} - \gamma_H^l) + p_H(\theta_{HM} - \gamma_M^h)$

[6]even though the aggregate surplus will depend on the actual

[7]when $(1-p_M)(\theta_{HM} - \gamma_M^l) + p_M(\theta_{HH} - \gamma_H^h) > (1-p_H)(\theta_{HH} - \gamma_H^l) + p_H(1-p_M)(\theta_{HM} - \gamma_M^l)$ and $(1-p_M)(\theta_{HM} - \omega_{FLSA}) + p_M(\theta_{HH} - \gamma_H^h) < (1-p_H)(\theta_{HH} - \gamma_H^l) + p_H(1-p_M)(\theta_{HM} - \omega_{FLSA})$

still make a sequential offer to group M. As a result, the substitution between high type and low type workers due to the overtime pay rule change can lead to inefficient situations where there is a lack of overtime labor supply.

The above two cases show that when there is substitution between workers of different level of productivity, the introduction of mandatory overtime pay could have different implications for overtime hours. It's possible for group H and M workers' overtime to either increase or decrease, depending on the substitution pattern. In addition, we consider the implications for overtime wage rate. Task $\bar{\theta}$ is exclusively reserved for high type workers. We assume that the productivity is high enough such that firms will always offer a high rate γ_H^h to make sure high type workers work. **Case 1** shows that it could increase the overtime hours of H group workers at compensation γ_H^l. Since the overtime wage for H group consists of γ_H^h from task $\bar{\theta}$ and γ_H^l from task θ_H, an increase in hours means the average rate for overtime hours will decrease. Likewise, if in **Case 2** the overtime hours for group H is reduced, we should expect to see an increase in average rate for overtime. Therefore, our model provides testable predictions: high type worker's change in overtime hours will be accompanied by a reverse change in overtime wage rate. The analysis for low type task θ_L will be very similar to the high type task: it's possible for group L type workers to either gain or lose overtime working hours when there is a mandatory overtime pay increase for M type worker.

2.3 Overtime and FLSA

2.3.1 Background information

The federal overtime provisions are specified in the Fair Labor Standard Act (FLSA). The Act, signed into law by President Franklin D. Roosevelt in 1938, was considered an important piece of New Deal legislation. The Act introduced several key regulations on the labor market, including 40-hour work week, national minimum wage,“time and a half" overtime pay. Over the past two decades, there has only been two majors changes to the overtime regulation. In 2004, the Bush administration revised the overtime rules regarding exemption criteria, raising the salary benchmark from \$250/week to \$445/week [8] and redefining job duties that qualify for exemption. It was the first major change on overtime rules since 1975. In 2016, the Department of Labor under the Obama administration published a final rule that updates the exemption criteria for workers. Slated to take effect on December 1st, 2016, the rule increased the salary threshold for exemption from \$455 to \$913 per week. However, on November 22, 2016, a federal judge blocked the rule, deciding that it is unlawful, and granted the motion for a nationwide injunction. As of today, the litigation regarding the rule is still in the appeal process, and with the new administration in charge, the rule is very much in jeopardy.

Overtime pay regulation has long been a controversial topic. Advocates calling for stricter overtime rule and higher overtime pay argue that it protects workers' rights and increase welfare for workers. Those against government regulation think that it would make labor more costly, hurting business, and they also question the effectiveness of the rules, since stricter overtime rules could make employers reduce working hours or baseline salary, ultimately having little or even negative impact on worker's income. When the new regulations on overtime were issued by the Department of Labor, they always came with specific policy goals and estimates of the impact on the labor market. However, opinions varied a lot regarding the actual impact of the rules. In 2004, the Department of Labor estimated that the updated rule would strengthen overtime protection for millions of salaried workers, transfer about \$375 million per year from employers to employees in the form of greater overtime pay or higher base salaries. The total first year implementation costs to employers are estimated to be \$738.5 million, and only an estimated 107000 workers who earn \$100000 or more per year could lose their overtime protection. [9] Nonetheless, many were critical of the new rule, saying that millions would lose eligibility of overtime pay because of the new definition of job duties: administrative workers can be reclassified as "team leaders", jobs such as cook or nurses can be categorized as "learned professional". The economic impact of overtime pay on workers was also heavily disputed, as some people suggested that baseline salary levels could

[8] There are two salary benchmarks in the previous rule, "long test" and "short test", with benchmark \$155/week and \$250/week, respectively.

[9] Economic Report, the Department of Labor, 29 CFR Part 541, 2004, https: www.dol.gov/whd/overtime/regulations.pdf

be re-adjusted to a lower level so that workers who work overtime would still earn the same income and work the same number of hours, rendering the new rule irrelevant.

Anecdotal examples on how the labor market adjusts to the overtime rule change are sporadic and paints a quite complicated picture. First of all, employers do not always passively follow the overtime rule as it is, they adjust to the rule change strategically. For example, in anticipation to the new overtime rule, which increased the exemption threshold from $455 to $913 per week, Wal-Mart raised the annual salaries for entry-level managers from $45000 to $48500 in order to avoid the unpredictable costs for salary employees, since many of them work overtime and the hours fluctuates from week to week.[10] Secondly, the overtime eligibility are often subject to the interpretation of the rule, and employers do not always enforce the rules strictly. In some firms, employees still get overtime pay even though they are exempted under the rule [11]; in some other cases, employers withhold overtime pay from workers and it's costly for workers to dispute, sometimes resulting in class-action lawsuits [12]. Sometimes, overtime-pay-exempted workers will still get compensated for weekly hours exceeding 40 due to tight labor market conditions[13].

2.3.2 2004 FLSA Rule Change

Most jobs in the United States are covered by the FLSA overtime provision, with few exceptions such as employees on foreign vessels, workers engaged in fishing operations or newspaper delivery, railroad and air carrier employees[14]. Workers covered by the overtime provision are either "Ex-

[10] Reuters News: http:
www.reuters.com/article/walmart-managers-overtime-idUSL1N1CH237

[11] The Record (New Jersey), August 15, 2004, "New Jersey employers confused about overtime eligibility as deadline nears"– "Of course, employers are free to continue paying overtime to workers who aren't covered by the new rules. And, according to a survey by Hewitt, about a third of employers pay overtime to workers who are not required by law to receive it."

[12] St. Petersburg Times (Florida), April 21, 2004, "Whitecollar OT rules are far from definitive" – "Pharmacists, current and former photo lab managers and assistant store managers of the Eckerd drugstore chain have filed class action lawsuits claiming the Largo company illegally denied them overtime."

[13] Las Vegas Review-Journal, August 24, 2004, "U.S. overtime rules to have little effect on Las Vegas payrolls, officials say" – " Registered nurses will no longer be entitled to overtime pay under the new law, but the shortage of nurses may cause many hospitals to keep paying overtime anyway. North Vista Hospital (formerly Lake Mead Hospital) will continue to pay overtime, human resources director Leanna Nalley said."

[14] For details, see "Handy Reference Guide to The Fair Labor Standards Act", Wage and Hour Division, Department of Labor, https:
www.dol.gov/whd/overtime/fs17a_overview.pdf

empted" or “Non-Exempted" for overtime pay. To be exempted, workers generally have to pass three tests: 1. Salary basis test (the employee has to be paid on a salary basis); 2. Salary level test (worker’s salary passes the designated threshold); 3. Duties test (workers need to have “white collar job" that require relatively high-level work, including executive, professional jobs)[15]. The 2004 FLSA revisions made several significant changes to the overtime rule that was last revised in 1975, including:

1. Increased minimum salary level from $250 to $455 per week.

2. Revised “job duties" tests replacing the current “long" and “short" tests with a single “standard duties" test for each exemption.

3. Highly compensated employees performing office or non-manual work whose annual salary is above $100000 are exempted from FLSA.

Under the revised rule, the group of workers directly affected are those “white collar workers" whose jobs satisfy the “Duties Test" description and have a salary between the old and new benchmark. Their status would change from “Exempt" to “Non-Exempt" and therefore they would become entitled to earn overtime compensation for work week that exceeds 40 hours. In our empirical study, we focus on two groups of workers: those whose overtime eligibility changed from “Exempt" to “Non-Exempt" because of the revision, and those whose status have always remained “Exempt". The main difference between these two groups is the salary level: the former has a salary level between the old and new benchmark before the rule change[16], the latter has a salary level above the new benchmark $455 per week before the rule change. We use worker’s occupation and income information to categorize their overtime eligibility. This approach is an approximation to worker’s exemption status since we do not have direct information on their overtime eligibility. To the extent that it classifies workers by their income and job, it allows us to study the effect on different groups of workers that the rule intended to treat differently.

One big controversy about the 2004 new overtime rule was that many thought it would lead to millions of workers losing overtime pay due to the redefined “Duties Test". The reasoning was that

[15] For details, see “Fact Sheet # 17A", Wage and Hour Division, Department of Labor, https: www.dol.gov/whd/overtime/fs17a_overview.pdf

[16] August 23rd, 2004

employers could reclassify workers into “Administrative", “Executive" or “Professional" employees so that they would pass the “Duties Test", especially since the description of duties had a lot of ambiguities and were subject to interpretation. However, there isn't much empirical evidence supporting this, and anecdotal evidences are quite mixed. By comparing detailed descriptions of job duties in the old and new overtime rules, there isn't any distinctive evidence that the rule made it easier for jobs to pass the “Duties Test"[17]. Furthermore, employers do not necessarily always follow the overtime rule: some would withhold overtime pay because litigation is costly for workers, some would pay for overtime hours, due to tight labor market supply, even though the employee is exempted. Given the complicated reality of overtime status, our approach allows us to study the aggregate effect on different worker groups. This classification method is also consistent with other empirical studies on this topic. For example, in Trejo (1991), detailed occupation classifications in the CPS were coded to determined the coverage status of each individual in the sample; in economic reports and impact studies issued by the Department of Labor [18] [19], information on worker's occupation were used to estimate the number of workers affected by overtime rules. To make sure classification errors are not altering the empirical results, we perform sensitivity tests by adding or removing different occupation categories into or from our group definitions, the results are not sensitive to these changes.

2.4 Empirical Models and Results

2.4.1 Data

Our data comes from two sources: Survey of Income and Program Participation (SIPP) and Current Population Survey (CPS). Both datasets offer unique advantages for studying this topic. CPS is a relatively more representative sample on the U.S. labor market and previous research mostly relies on it; SIPP offers a longer panel at monthly level and contains good quality information on worker's

[17] The Department of Labor listed a comparison of duties tests on the website https: www.dol.gov/whd/regs/compliance/overtime/side-by-side_PF.htm

[18] Economic Report, the Department of Labor, 29 CFR Part 541, 2004, https: www.dol.gov/whd/overtime/regulations.pdf

[19] Summary of the Economic Impact Study, the Department of Labor, https: www.dol.gov/whd/overtime/final2016/overtimeFinalRule.pdf

hours.

We use the 2004 SIPP panel dataset[20] to study the effect of overtime rule change. Set D of the panel contains data for analysis on labor force participation, employment, and earnings. The panel spans from July 2003 to Dec 2008, therefore it allows us to have observations on workers' income and working hours before and after the rule change.

For CPS, the panel is constructed by merging survey data on the CPS Outgoing Rotation Groups. Every household in the CPS sample was interviewed consecutively for 4 months, then left out for 8 months, and interviewed again for 4 more months. Weekly hours and earning questions are asked on their 4th and 8th interview. As new households enter into the survey each month, one fourth of the households are in the outgoing rotation each month. Therefore, we have a balanced two-period panel that spans exactly one year in-between observations. To study the effect of rule change, we choose job spells with observations both before and after the rule change. In comparison to SIPP which usually has 13 months of observation for each person, the CPS panel is much shorter in time series.

2.4.2 Overtime Exemption Status Classification

The objective of categorizing workers is to study the impact on workers with different overtime pay statue under the rule. We focus on two groups of workers. First, we are interested in those who were directly impacted by the rule due to the increased salary benchmark. These are salary workers whose jobs fit the descriptions of the duties test and earned salary between the old and new benchmark. A key issue about overtime law is how effective mandatory overtime pay is at improving worker's welfare, since employers have many options to offset the intended policy outcome. The second group we are interested in are workers who remained exempted from overtime pay during the rule change period.

We classify workers mainly based on their income and occupation. Both SIPP and CPS data have detailed occupation information based on four-digit Census Occupation Codes [21]. First, we identify the group of workers that are exempted from overtime pay directly. They are either ex-

[20]We thank Center for Economic and Policy Research for providing Uniform Extracts of SIPP datasets.

[21]For a complete list, see https:
cps.ipums.org/cps/codes/occ_20032010_codes.shtml

empted from all FLSA coverages or only the overtime provision. Examples include farmers employed on small farms, railroad and air carrier employees. Second, we label occupations that would most likely pass the "Duties Test" based on the Department of Labor's duty description. These occupations consists of four major categories: "Executive Employees", "Administrative Employees", "Professional Employees" and "Computer Employees". We leave out "Outside Sales Employees" because it's hard to decide if workers in sales-related occupations satisfy the duties description, and there wasn't a salary level requirement before or after the rule change, therefore it's hard to tell if they are directly affected by the rule change. In Appendix A, we list the occupation codes under each categories. Then we classify workers by their income level. We definite the group of workers as "Between Group" if their salary falls into the old and new salary benchmark interval. For simplicity, we use the "short test" benchmark ($250 per week) in the old rule. It is consistent with the impact studies issued by the Department of Labor. Moreover, by choosing the relatively higher benchmark (the "long test" has a benchmark $155 per week), we identify the group of workers with higher probability to gain eligibility of overtime pay, since workers with weekly salary between $155 and $250 could be eligible for overtime pay before the rule change using the "long test". In our robustness check, we shift the salary benchmark to different levels and the results are very similar. We also exclude those who earn more than $100000 in annual income, since they are not covered by FLSA overtime provision and usually top-coded in the data. In the SIPP panel, information on income and hours is at monthly level, we use average weekly income to label income group. CPS outgoing rotation group data, however, has information on hours worked in the week prior to the interview. Our empirical analysis is at monthly level for SIPP, and is at weekly level for CPS.

We divide jobs in our sample into two categories based on labor income: salary or hourly wage. These two categories are then further divided based on their overtime exemption status. Wage jobs in general are eligible for overtime pay, except for a few occupations not covered by FLSA overtime rule. We label these jobs as "Exempt" under the wage job category, and label the rest as "Non-Exempt"[22]. For SIPP, we have less than 1% out of all the wage job observations that are exempted based on this. The categorization for salary jobs is a bit more detailed. To be exempt from

[22]For example: clergy, railroad conductors, captains and other officers on fishing vessels, see Appendix A

overtime pay, salary workers have to pass both the "Salary Level Test" and the "Duties Test". We focus on three subgroups within salary groups. "Treated" group refers to jobs that would pass the "Duties Test" whose income falls in the "Between Group". "Exempt" group refers to jobs that had salary above the $455 new benchmark and pass the "Duties Test". These are mainly high income administrative, executive or professional jobs. The "Exempt" group in the sample mainly serves as comparison to the "Treated' group since their job duties are similar. We also pick "Non-Exempt" group that has similar income level as a reference group. For this group, the income level also falls in the "Between Group", but workers' occupations most likely wouldn't pass the "Duties Test" and therefore they would not be eligible for overtime pay after the rule change. "Low Income" group refers to the group of workers whose income was below the original salary test threshold $250 per week. We use only the income criteria to select "Low Income" group in order to have reasonable number of job spells. In our robustness check, we narrow the definition by restricting to job spells that pass the "Duties Test". Such a group classification allows us to identify the effects of mandatory overtime pay for various groups and test the predictions from our adverse selection model.

2.4.3 The Effect of Overtime Rule Change on Hours and Income

We use the following model to study the direct impact of rule change on hours and income:

$$
\begin{aligned}
Y_{it} = \alpha_0 Treated_i * RuleChange_t + \alpha_1 Exempt_i * RuleChange_t \\
+ \alpha_2 LowIncome_i * RuleChange_t + X_{it}\beta + \delta_t + \xi_i + \varepsilon_{it}
\end{aligned}
$$

$Treated_i$ is the group dummy for workers that gained overtime eligibility under the new rule. $Exempt_i$ and $LowIncome_i$ are defined similarly. $RuleChange_t$ is a time dummy for observations after the rule change[23]. We include job and time fixed effects in our model. X_{it} includes worker level control variables: age, education, martial and union status. The coefficients α_0, α_1 and α_2 are the effects of rule change on "Treated", "Exempt" and "LowIncome" groups identified by diff-in-diff: the

[23]The rule change was effective starting August 23rd, 2004, we set September 2004 and afterwards as post rule change period.

difference between the change of each group before and after the rule change. We use this model to study the impact of rule change on working hours, overtime working and wage. Since salary jobs are directly affected by overtime rules and quite different from wage jobs in many aspects, we exclude wage jobs from the sample when estimating the model.

Table 2.3 shows the baseline results from SIPP panel. For the “Treated" group, the rule is estimated to increase working hours for about 2%, and increase wage rate and monthly income for about 15%. The increase in working hours, wage rate and total income for the “Treated" group is consistent with our adverse selection model for overtime labor: mandatory overtime pay provides a “price floor" in an otherwise monopolistic labor market for the firm, as a result, workers with higher preference for leisure will be sufficiently compensated to supply overtime work, resulting in increases in overtime working, higher wage rate and total labor income. The coefficients on “Exempt" and “LowIncome" also lends support to our multi-segment model for overtime working. There is a significant decline in overtime working for the “Exempt" group after the rule change, while the coefficient sign for hourly wage rate and total labor income is positive. This corresponds to **Case 2** in the multi-segment model, where firms prioritize “high" type tasks to group H worker but exclude the high leisure group γ_H^h: overtime work for group H decreases, but the average wage rate increases, since the proportion of highly compensated task $\bar{\theta}$ among overall overtime work has increased.

We also estimate the same model using CPS panel. Table 2.4 shows a larger effect on income: hourly rate and weekly income increased for about 36%, but does not show any meaningful impact on the hours worked. Similar to SIPP panel, the effect on “Exempt" group is quite insignificant. We use different sample scopes of CPS in estimation and the results are quite robust. We also include “Non-Exempt" group in one of our model specifications and find a 23% increase in income after the rule change, less than the “Treated" group but more than the “Exempt" group.

To further examine if the results are driven by a general time trend or a sudden change in a particular year, we estimate the Diff-in-Diff coefficients by year:

$$Y_{it} = \alpha_0 Treated_i * RuleChange_t * YearDummy_t + \alpha_1 Exempt_i * RuleChange_t * YearDummy_t$$
$$+ X_{it}\beta + \delta_t + \xi_i + \varepsilon_{it}$$

A similar model is estimated for CPS, but using monthly dummy instead. Table 2.5 shows that the rule change impact were relatively stable across years. For example, hourly wage rate increase is estimated at about 15% from 2005 to 2007 each year, except for 2004, which showed a lower magnitude at about 8%. But it should also be noted that the rule change was officially enacted on late April 2004.

To compare the group difference before and after rule change, we plot residuals from baseline regression $Y_{it} = X_{it}\beta + \delta_t + \xi_i + \varepsilon_{it}$ and estimated monthly Diff-in-Diff time effects over time.Figure 2.1 and 2.2 shows the graphs for weekly hour and income. For each variable of interest, we plot three graphs:

1. the residual from baseline regression, plotted over time. To simplify the visual, we plot different percentiles of residual distribution and trace the distribution change over time by mapping fitted time series graph over time.

2-3. Diff-in-Diff estimates of "Treated" group against baseline control group. We plot the "Treated" group time effect together with time effects for "Exempt" and "Non-Exempt" group, in the second and third graph respectively for each key variable.

The graphs from SIPP results illustrate the identification of rule change effect: for "Treated" group, we see a notable increase in income and wage after the rule change and the magnitude of increase remain relatively stable. In comparison, "Exempt" and "Non-Exempt" groups remain relatively flat. This shows that our estimation results are not driven by a general time trend in the group difference.

Similarly, Figure 2.3 plots the group difference over the rule change period for estimates from CPS panel. The results are a bit different from those in the SIPP panel: consist with estimation results for working hours, there isn't much significant difference between groups overtime, the estimated effects of rule change on wage and income seem to be a general group difference, since even before the rule change we also observe a relative increase in income for treated group that is of similar magnitude.

Overall, we find some evidence that the 2004 FLSA rule change increased the wage and income for salaried workers who were directly impacted by the increase in salary level test threshold, but the results are a bit mixed across different datasets.

2.4.4 Income and Wage models

Under FLSA, workers who are not exempted from overtime pay would receive "time-and-half" pay if their work exceeds 40 hours per week. To study the real effect of the rule change on overtime pay premium, we need to compare the relative rate of overtime hours against regular hourly rate before and after the rule change. In the data, we rarely have good measure of the overtime pay workers receive for a given period of time. In addition, it's not unusual for workers to get some extra labor income in the form of performance pay, such as commission or bonus. To circle around this problem, we use the following model to derive overtime pay premium:

Let I_t be the total labor income for a worker in time period t, and S_t be the salary. Given that a worker works for h_t hours in the period, his effective average wage rate is $w_t = I_t/h_t$. Let S be the worker's base salary, his normal hourly rate is $S/\bar{h}$ where $\bar{h}$ is the threshold for overtime work (40 hours per week). The total labor income is then modeled as the following:

$$I_t = w_t h_t = S(1 + \beta \frac{h - \bar{h}}{\bar{h}})$$

Taking logarithm of both sides, we decompose labor income into salary part and overtime pay:

$$\log I_{it} = \log S + \log(1 + \beta \frac{h_{it} - \bar{h}}{\bar{h}}) \approx \log S + \beta \frac{h_{it} - \bar{h}}{\bar{h}}$$

based on the equation above, we derive the following empirical models:

$$\log I_{it} = \beta_0 \frac{h_{it} - \bar{h}}{\bar{h}} * RuleChange_t + \beta_1 \frac{h_{it} - \bar{h}}{\bar{h}} + X_{it}\gamma + \delta_t + \xi_i + \varepsilon_{it}$$

or, equivalently,

$$\log w_{it} = -\alpha \log(h_{it}) + \beta_0 \frac{h_{it} - \bar{h}}{\bar{h}} * RuleChange_t + \beta_1 \frac{h_{it} - \bar{h}}{\bar{h}} + X_{it}\gamma + \delta_t + \xi_i + \varepsilon_{it}$$

With job and time fixed effects, we are able to get round the baseline salary for the job since usually we only observe the total labor income that includes overtime pay. The coefficients for variables containing $\frac{h_{it}-\bar{h}}{\bar{h}}$ measure the relative rate at which overtime hours are compensated for. β_0 is then the identified effect of rule change on overtime pay rate.

We estimate a pooled-sample model in which we include all the salary job samples in the model and use group dummies to differentiate the impact on workers with different exemption status:

$$\begin{aligned}\log I_{it} = {} & \beta_0 GroupDummy_i * RuleChange_t * \frac{h_{it} - \bar{h}}{\bar{h}} + \alpha_1 GroupDummy_i * \frac{h_{it} - \bar{h}}{\bar{h}} \\ & + \alpha_2 GroupDummy_i * RuleChange_t + \alpha_3 RuleChange_t * \frac{h_{it} - \bar{h}}{\bar{h}} \\ & + \beta \frac{h_{it} - \bar{h}}{\bar{h}} + X_{it}\gamma + \delta_t + \xi_i + \varepsilon_{it}\end{aligned}$$

As Table 2.6 and 2.7 shows, the average effect of rule change on overtime pay premium is about 0.27 for the "Treated" group. This means that overtime rates increased by about 27% over the normal rate after the rule change. The estimated effect has a quite large standard error (0.24) and is not significant. If the rule change had the intended outcome for workers who gained overtime eligibility, we should expect to see a significant estimated effect on the "Treated" group around 0.5. There are a few possible explanations for this. First, as Trejo (1991) pointed out in the "Fixed-Job Model", employers could lower the normal wage rate in response to the rule, this would lower the overtime pay premium estimated from data. Moreover, as anecdotal evidence suggests, overtime pay is not always used exactly as required by the law. This means employers could withhold overtime pay even when the employee is eligible for it. Another interesting finding is that there was a small but significant increase in overtime pay premium for "Exempt" group. The estimated coefficient 0.076 means that after the rule change the overtime rate has increase by about 7% over the normal rate. There isn't any direct mechanism from the rule change to increase the overtime premium for exempted workers. One possible explanation is the externality from the “Treated" group, as some workers in the same firm are getting extra compensation, it's hard to leave out other employees, and updating the payroll system could be costly. It could also be a spill-over effect due to reduced labor supply, as our adverse selection model on multi-segments of workers suggests, “Exempt" group workers with higher preference for leisure could be priced out of the market, leaving only the high paying overtime job for them and hence increase the average overtime pay premium.

Another interesting finding is the group difference in overtime pay premium. The coefficients for $Treated_i * \frac{h_{it}-\bar{h}}{\bar{h}}$ and $Exempt_i * \frac{h_{it}-\bar{h}}{\bar{h}}$ are -0.55 and -0.04 respectively. This means before the rule change, the average overtime premium is significantly lower for these two groups compared with the rest of the salary jobs. The magnitude of difference is much larger for the “Treated" group, it implies a 55% less overtime premium than other salary jobs. This is consistent with “time-and-half" overtime pay, since “Treated" group workers were not exempted from overtime pay before the rule change.

A counter-intuitive result is the estimated magnitude of overtime pay rate. As in our baseline model, β is the relative rate of overtime pay with respect to regular hourly rate. If salary workers get compensated for extra hours, the rate should be at least close to the regular hours. In the income and overtime pay premium model (Table 2.7), for salary jobs in our sample, we get very small estimates of β for both exempt and non-exempt workers, 0.011and 0.002 respectively, and the estimates for “Treated" group is as low as as -0.516 (this however, is consistent with the “time-and-half" difference as pointed out earlier). There are two possible explanations for this. First, there might be a prevalent non-compliance of the FLSA even though the "Non-Exempt" jobs are statutorily qualified for overtime pay. Second, the wage and income models do not account for non-hour related variable labor compensation such as bonus: if workers could get the task done within regular hours and get rewarded for high quality work, we will then observe a very low rate of overtime compensation because the non-hour-related performance pay drives up baseline salary in our model.

Overall, we find an increase in overtime pay premium after the rule change for workers who gained overtime eligibility under the new rule. There was on average a 26% increase with respect to normal rate. However, the effect varies across jobs in the group and is not significant. Meanwhile, there is a small but significant increase in overtime premium for “Exempt" group. Their overtime premium rate increased by about 8% after the rule change. This might be due to the externality from the directly impacted group on labor market supply, as predicted in our adverse selection model.

2.4.5 Robustness of Results

To test if our results are sensitive to the group identification methods, we change the criteria and re-estimate the models to see if the estimates are robust to slight change of definition.

We first test if the results are sensitive to the time frame of our income group definition. One drawback of the income group identification is that it “forces" the income level to be between the old and new salary benchmark. As we examine the income and wage rate after the rule change, it's possible to see an increase in them just because of this selection issue: income will increase for that group because the group selection are only based on income level before the rule change. To address this concern, we use a placebo test to estimate the effect. We choose June 2004 as the “placebo" rule timing - 3 month prior to the actual rule change. Similar to the previous analysis, we plot the month-by-month estimates to show the trend. Figure 2.4 plots the time trend of the coefficients from the placebo test, for the effects on hours and income as well as the wage model. The first graph shows the Diff-in-Diff estimates for overtime wage premium. There is a sizable increase in coefficients after the actual rule change, this suggests that the overtime premium only starts to increase for the “Treated" group against the rest of salary jobs when the new rule is implemented. The second and third graph plots the estimated trend for wage and weekly working hours. The wage increase is stabilized after the actual rule change, there is an increasing trend that started 2 months prior to the actual rule change, but compared with the original estimates in Figure 2.2, the trend and timing of the increase are still very similar. The trend of weekly working hours is also similar to the original estimates. The placebo test reaffirms that our results are robust to the timing of income group selection criteria.

A second concern is the choice of income benchmark. The old rule has a "Long Test" which requires only a weekly income of \$155. In addition, since our weekly income is imputed using total income and number of weeks worked in the month, we might overestimate workers' salary basis since they could receive performance based pay in the form of tips, commission or bonus. Therefore, we relax the income benchmark to see if our results are driven by a sub-group of workers. Table 2.8 shows the wage and overtime pay premium model estimation. The results are quite similar to our previous findings: the overtime pay premium increased by about 22.1%for the "Treated" group and is significant at 10% level. We also get similar coefficients for the "Exempt" group. Table 2.9 shows the estimated effects on hours and income. Weekly hours increased by 1.9% and wage increased by 6.6%. The estimates are slightly lower than our original estimates because the baseline selection criteria is more restrictive. By relaxing the income selection range, we are more likely to include jobs that were not directly affected by the rule change.

2.5 Conclusion

We study the effect of the revision to statutory overtime pay under Fair Labor Standard Acts in 2004 and find evidence that workers' hours and income increased for those who gained statutory overtime pay coverage. We also find spill-over effects on the overtime pay premium and overtime schedule on workers who are not directly affected by the rule change. Our results suggest that overtime pay regulation does have tangible effects on the labor market. Contrary to the belief that firms can simply adjust the wage rate in response to the mandatory overtime pay premium such that the working schedule and compensation remain de-facto unaffected, our results suggest that workers benefit from increased overtime pay. The result is consistent with our adverse selection model and suggests that overtime pay regulation has effects on the contractual and informational aspects of the supply and demand for overtime work. Further research on the contract aspect for overtime pay and the mechanism through which it affects both the supply and demand of labor would be interesting.

	Without FLSA			With FLSA			FLSA effect	
Firm demand and worker supply	OT pay	High type	Low type	OT pay	High type	Low type	Hours	Efficiency
$\gamma_L < \theta < \gamma_H$	γ_L	N	Y	$\gamma_L < \omega_{FLSA} < \theta$	N	Y	Same	Same
				$\gamma_L < \theta < \omega_{FLSA}$	N	N	Decrease	Decrease
$\gamma_L < \gamma_H < \theta$	γ_H	Y	Y	$\gamma_H < \theta < \omega_{FLSA}$	N	N	Decrease	Decrease
$(p_\gamma \theta > \gamma_H - (1 - p_\gamma)\gamma_L)$				$\gamma_H < \omega_{FLSA} < \theta$	Y	Y	Same	Same
				$\gamma_L < \omega_{FLSA} < \gamma_H$	Y	Y	Same	Same
$\gamma_L < \gamma_H < \theta$	γ_L	N	Y	$\gamma_H < \theta < \omega_{FLSA}$	N	N	Decrease	Decrease
$(p_\gamma \theta < \gamma_H - (1 - p_\gamma)\gamma_L)$				$\bar{\gamma}^{(*)} < \omega_{FLSA} < \theta$	Y	Y	Increase	Increase
				$\gamma_L < \omega_{FLSA} < \bar{\gamma}^{(*)}$	N	Y	Same	Same

*The cutoff $\bar{\gamma} = \frac{\gamma_H - p_\gamma \theta}{1 - p_\gamma}$. When mandatory overtime pay ω_{FLSA} is above this level, firms will find it more profitable to raise the pay to $max(\gamma_H, \omega_{FLSA})$ to induce all workers working overtime.

Initial Target	Offer	Firm's expected profit	Overtime Probability H	M
H	γ_H^h	$\theta_{HH} - \gamma_H^h$	1	0
H	(γ_H^l, γ_M^h)	$(1-p_H)(\theta_{HH} - \gamma_H^l) + p_H(\theta_{HM} - \gamma_M^h)$	$(1-p_H)$	p_H
H	(γ_H^l, γ_M^l)	$(1-p_H)(\theta_{HH} - \gamma_H^l) + p_H(1-p_M)(\theta_{HM} - \gamma_M^l)$	$(1-p_H)$	$p_H(1-p_M)$
M	γ_M^h	$\theta_{HM} - \gamma_M^h$	0	1
M	(γ_M^l, γ_H^h)	$(1-p_M)(\theta_{HM} - \gamma_M^l) + p_M(\theta_{HH} - \gamma_H^h)$	p_M	$(1-p_M)$
M	(γ_M^l, γ_H^l)	$(1-p_M)(\theta_{HM} - \gamma_M^l) + p_M(1-p_H)(\theta_{HH} - \gamma_H^l)$	$p_M(1-p_H)$	$(1-p_M)$

Table 2.1: Task θ_H delegation and corresponding payoffs

Initial Target	Offer	Firm's expected profit	Overtime Probability H	M
H	γ_H^h	$\theta_{HH} - \gamma_H^h$	1	0
H	(γ_H^l, γ_M^h)	$(1-p_H)(\theta_{HH} - \gamma_H^l) + p_H(\theta_{HM} - \gamma_M^h)$	$(1-p_H)$	p_H
H	$(\gamma_H^l, \omega_{FLSA})$	$(1-p_H)(\theta_{HH} - \gamma_H^l) + p_H(1-p_M)(\theta_{HM} - \omega_{FLSA})$	$(1-p_H)$	$p_H(1-p_M)$
M	γ_M^h	$\theta_{HM} - \gamma_M^h$	0	1
M	$(\omega_{FLSA}, \gamma_H^h)$	$(1-p_M)(\theta_{HM} - \omega_{FLSA}) + p_M(\theta_{HH} - \gamma_H^h)$	p_M	$(1-p_M)$
M	$(\omega_{FLSA}, \gamma_H^l)$	$(1-p_M)(\theta_{HM} - \omega_{FLSA}) + p_M(1-p_H)(\theta_{HH} - \gamma_H^l)$	$p_M(1-p_H)$	$(1-p_M)$

Table 2.2: Task θ_H delegation and corresponding payoffs with FLSA Overtime rule change

Dependent Variable	log hours (monthly)	log hours (weekly)	Overtime Dummy	log income (monthly)	log wage (hourly)	wage (hourly)
Treated*Rule	0.025*	0.025*	0.013	0.162***	0.138***	1.940***
	(0.01)	(0.01)	(0.02)	(0.04)	(0.04)	(0.69)
Exempt*Rule	-0.002	-0.002	-0.016**	0.003	0.005	1.387***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.26)
LowIncome*Rule	0.007	0.007	0.007	0.048	0.041	0.361
	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.49)
Age	0.008***	0.008***	0.010*	0.018***	0.011*	-0.196
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.25)
$Age^2/100$	-0.009***	-0.009***	-0.013**	-0.022***	-0.013*	0.314
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.28)
Edu	-0.001	-0.001	-0.007	0.007	0.008	0.341
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.31)
Union	0.006	0.006	0.002	0.033	0.027	0.705
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.56)
Married	0.001	0.001	0.002	0.000	-0.001	0.006
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.09)
R^2	0.005	0.005	0.007	0.025	0.026	0.008
N	358077	358077	358077	355842	355842	355842

Table 2.3: Effect of rule change on hours and income

1. **Model:** $Y_{it} = \alpha_0 Treated_i * Rule_t + \alpha_1 Exempt_i * Rule_t + \alpha_2 LowIncome_i * Rule_t + X_{it}\beta + \delta_t + \xi_i + \varepsilon_{it}$
2. RuleChange is time dummy for enactment of 2004-FLSA regulations change. (2004 September)
3. 2004 SIPP Panel, from year 2003 (starting July) to 2007
4. Samples exclude those who work for different jobs at the same time, or self employed
5. “Treated" group refers to job spells that are exempted before 2004 rule change and not exempted after
6. Standard Errors clustered at job spell level
7. Year-Month and state fixed effect included

Dependent Variable	log hours (weekly)	Overtime Dummy	log income (weekly)	log wage (hourly)	Extra payment	Get overtime
Only samples work on same job						
Treated*rule	0.010	0.002	0.366***	0.362***	-0.009	0.001
	(0.02)	(0.03)	(0.05)	(0.06)	(0.01)	(0.01)
Exempt*rule	0.003	0.015	-0.000	-0.010	-0.026***	-0.001
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
r2	0.004	0.002	0.017	0.022	0.038	0.006
N	16722	16726	16673	15666	16726	16726
Only samples work on same job						
Treated*rule	0.012	0.004	0.389***	0.384***	-0.008	0.001
	(0.02)	(0.03)	(0.05)	(0.06)	(0.01)	(0.01)
Exempt*rule	0.005	0.017	0.022*	0.012	-0.025***	-0.000
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Non-Exempt*rule	0.012	0.011	0.198***	0.193***	0.009	0.004
	(0.01)	(0.02)	(0.03)	(0.03)	(0.01)	(0.00)
r2	0.004	0.002	0.023	0.029	0.038	0.006
N	16722	16726	16673	15666	16726	16726
All samples include those who switch jobs						
Treated*rule	0.015	0.008	0.438***	0.425***	-0.005	-0.000
	(0.01)	(0.02)	(0.03)	(0.03)	(0.01)	(0.00)
Exempt*rule	0.008***	0.014**	0.037***	0.027***	-0.021***	-0.001
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Non-Exempt*rule	0.007	0.026**	0.232***	0.239***	0.008	0.004
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.00)
r2	0.003	0.003	0.022	0.028	0.041	0.009
N	37372	37382	37248	35252	37382	37382
Wage jobs						
Rule	0.009	0.000	0.051**	0.024	-0.036**	0.001
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
r2	0.001	0.002	0.010	0.008	0.012	0.002
N	20785	20794	20772	20671	20794	20794

Table 2.4: Effect of rule change on hours and income (CPS)

1. **Model:** $Y_{it} = \alpha_0 Treated_i * RuleChange_t + \alpha_1 Exempt_i * RuleChange_t + X_{it}\beta + \delta_t + \xi_i + \varepsilon_{it}$
2. RuleChange is time dummy for enactment of 2004-FLSA regulations change. (2004 September)
3. CPS Panel
4. Samples exclude those who work for different jobs at the same time, or self employed
5. "Treated" group refers to jobs spells that are exempted before 2004 rule and not exempted after
6. Standard Errors clustered at job spell level
7. Year-Month fixed effect included
8. Extra payment is the dummy for if worker usually gets overtime/tips/commission.
9. Get overtime is the dummy for to if worker work overtime and paid extra

Dependent Variable	log hours (monthly)	log hours (weekly)	Overtime Dummy	log income (monthly)	log wage (hourly)	wage (hourly)
Treated*Rule*2004	0.003	0.003	0.006	0.086***	0.084**	1.467***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.57)
Treated*Rule*2005	0.029**	0.029**	0.013	0.170***	0.143***	1.980***
	(0.01)	(0.01)	(0.02)	(0.04)	(0.04)	(0.67)
Treated*Rule*2006	0.027**	0.027**	0.024	0.187***	0.161***	2.022**
	(0.01)	(0.01)	(0.03)	(0.05)	(0.05)	(0.87)
Treated*Rule*2007	0.047*	0.047*	-0.004	0.217***	0.172**	2.559
	(0.03)	(0.03)	(0.05)	(0.07)	(0.08)	(1.66)
Exempt*Rule*2004	-0.003	-0.003	-0.021***	0.002	0.005	0.895***
	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.27)
Exempt*Rule*2005	-0.002	-0.002	-0.016**	0.001	0.003	1.304***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.29)
Exempt*Rule*2006	-0.003	-0.003	-0.010	0.005	0.007	1.934***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.38)
Exempt*Rule*2007	0.002	0.002	-0.017	-0.004	-0.006	1.533***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.56)
Age	0.008***	0.008***	0.010*	0.018***	0.010	-0.205
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.25)
$Age^2/100$	-0.009***	-0.009***	-0.013**	-0.022***	-0.013*	0.324
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.28)
Edu	-0.001	-0.001	-0.007	0.008	0.008	0.346
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.31)
Union	0.006	0.006	0.002	0.033	0.027	0.701
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.56)
Married	0.001	0.001	0.002	0.000	-0.001	0.005
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.09)
r2	0.005	0.005	0.008	0.025	0.026	0.008
N	358077	358077	358077	355842	355842	355842

Table 2.5: Effect of rule change on hours and income (by year)

1. **Model:** $Y_{it} = \alpha_0 Treated_i * RuleChange_t * YearDummy_t + \alpha_1 Exempt_i * RuleChange_t * YearDummy_t + X_{it}\beta + \delta_t + \xi_i + \varepsilon_{it}$
2. RuleChange is time dummy for enactment of 2004-FLSA regulations change. (2004 September)
3. 2004 SIPP Panel, from year 2003 (starting July) to 2007
4. Samples exclude those who work for different jobs at the same time, or self employed
5. “Treated" group refers to jobs spells that are exempted before 2004 rule and not exempted after
6. Standard Errors clustered at job spell level
7. Year-Month and state fixed effect included

Dependent Variable	log wage (hourly)							
	Salary Jobs						Wage Jobs	
	All	All	Treated	Non-Exempt	LowIncome	Exempt	Non-Exempt	Exempt
Treated$\times Rule \times \frac{h_{it}-\bar{h}}{\bar{h}}$	0.272							
	(0.24)							
Exempt$\times Rule \times \frac{h_{it}-\bar{h}}{\bar{h}}$	0.076***							
	(0.02)							
LowIncome$\times Rule \times \frac{h_{it}-\bar{h}}{\bar{h}}$	0.138							
	(0.33)							
Treated$\times\frac{h_{it}-\bar{h}}{\bar{h}}$	-0.562**							
	(0.26)							
Exempt$\times\frac{h_{it}-\bar{h}}{\bar{h}}$	-0.046**							
	(0.02)							
LowIncome$\times\frac{h_{it}-\bar{h}}{\bar{h}}$	-0.314							
	(0.26)							
Treated$\times Rule$	0.137***							
	(0.03)							
Exempt$\times Rule$	-0.008							
	(0.01)							
LowIncome$\times Rule$	0.038							
	(0.03)							
Rule=1$\times\frac{h_{it}-\bar{h}}{\bar{h}}$		0.047***	0.268	0.106	0.148	0.078***	-0.001	0.110
		(0.01)	(0.24)	(0.08)	(0.33)	(0.02)	(0.02)	(0.20)
$\frac{h_{it}-\bar{h}}{\bar{h}}$	-0.085***	-0.121***	-0.563**	-0.229**	-0.375	-0.100***	-0.103***	-0.398
	(0.02)	(0.02)	(0.25)	(0.10)	(0.26)	(0.03)	(0.02)	(0.26)
$\log(h_{it})$	-0.824***	-0.823***	-0.944***	-0.698***	-0.863***	-0.864***	-0.047***	-0.096
	(0.02)	(0.02)	(0.06)	(0.10)	(0.07)	(0.04)	(0.01)	(0.07)
R^2	0.191	0.190	0.372	0.159	0.365	0.245	0.036	0.084
N	355842	355842	3730	11232	4600	123338	575263	3654

Table 2.6: Wage and overtime premium Model

1. **Model:** $\log w_{it} = -\alpha \log(h_{it}) + \beta_0 \frac{h_{it}-\bar{h}}{\bar{h}} * RuleChange_t + \beta_1 \frac{h_{it}-\bar{h}}{\bar{h}} + X_{it}\gamma + \delta_t + \xi_i + \varepsilon_{it}$
2. RuleChange is time dummy for enactment of 2004-FLSA regulations change. (2004 September)
3. “Treated" group refers to jobs spells that are exempted before 2004 rule and not exempted after
4. “LowIncome" group refers to job spells that have lower income level than “Treated" group prior to rule change

Dependent Variable	log wage (hourly)							
	Salary Jobs						Wage Jobs	
	All	All	Treated	Non-Exempt	LowIncome	Exempt	Non-Exempt	Exempt
Treated$\times Rule \times \frac{h_{it}-\bar{h}}{\bar{h}}$	0.265							
	(0.24)							
Exempt$\times Rule \times \frac{h_{it}-\bar{h}}{\bar{h}}$	0.078***							
	(0.02)							
LowIncome$\times Rule \times \frac{h_{it}-\bar{h}}{\bar{h}}$	0.137							
	(0.33)							
Treated$\times \frac{h_{it}-\bar{h}}{\bar{h}}$	-0.554**							
	(0.26)							
Exempt$\times \frac{h_{it}-\bar{h}}{\bar{h}}$	-0.045**							
	(0.02)							
LowIncome$\times \frac{h_{it}-\bar{h}}{\bar{h}}$	-0.303							
	(0.27)							
Treated$\times Rule$	0.141***							
	(0.03)							
Exempt$\times Rule$	-0.008							
	(0.01)							
LowIncome$\times Rule$	0.039							
	(0.03)							
Rule=1$\times \frac{h_{it}-\bar{h}}{\bar{h}}$		0.049***	0.266	0.111	0.148	0.079***	0.047**	-0.081
		(0.01)	(0.24)	(0.09)	(0.34)	(0.02)	(0.02)	(0.23)
$\frac{h_{it}-\bar{h}}{\bar{h}}$	0.058***	0.022	-0.516**	0.002	-0.254	0.011	0.332***	0.231
	(0.01)	(0.01)	(0.25)	(0.07)	(0.27)	(0.02)	(0.02)	(0.21)
R^2	0.026	0.025	0.187	0.043	0.084	0.044	0.028	0.061
N	355842	355842	3730	11232	4600	123338	575263	3654

Table 2.7: Income and overtime premium Model

1. **Model:** $\log I_{it} = \beta_0 \frac{h_{it}-\bar{h}}{\bar{h}} * RuleChange_t + \beta_1 \frac{h_{it}-\bar{h}}{\bar{h}} + X_{it}\gamma + \delta_t + \xi_i + \varepsilon_{it}$
2. RuleChange is time dummy for enactment of 2004-FLSA regulations change. (2004 September)
3. “Treated" group refers to jobs spells that are exempted before 2004 rule and not exempted after
4. “LowIncome" group refers to job spells that have lower income level than “Treated" group prior to rule change

Dependent Variable	log wage (hourly)						
	Salary Jobs					Wage Jobs	
	All	All	Treated	Non-Exempt	Exempt	Non-Exempt	Exempt
Treated$\times Rule \times \frac{h_{it}-\bar{h}}{\bar{h}}$	0.221*						
	(0.12)						
Exempt$\times Rule \times \frac{h_{it}-\bar{h}}{\bar{h}}$	0.076***						
	(0.02)						
Treated$\times \frac{h_{it}-\bar{h}}{\bar{h}}$	-0.241*						
	(0.14)						
Exempt$\times \frac{h_{it}-\bar{h}}{\bar{h}}$	-0.045**						
	(0.02)						
Treated$\times Rule$	0.070***						
	(0.02)						
Exempt$\times Rule$	-0.006						
	(0.01)						
Rule=1$\times \frac{h_{it}-\bar{h}}{\bar{h}}$		0.047***	0.216*	0.117*	0.078***	-0.001	0.110
		(0.01)	(0.12)	(0.07)	(0.02)	(0.02)	(0.20)
$\frac{h_{it}-\bar{h}}{\bar{h}}$	-0.090***	-0.121***	-0.292*	-0.118*	-0.097***	-0.103***	-0.398
	(0.02)	(0.02)	(0.15)	(0.07)	(0.03)	(0.02)	(0.26)
$\log(h_{it})$	-0.824***	-0.823***	-0.871***	-0.814***	-0.871***	-0.047***	-0.096
	(0.02)	(0.02)	(0.08)	(0.05)	(0.04)	(0.01)	(0.07)
r2	0.191	0.190	0.295	0.225	0.248	0.036	0.084
N	355842	355842	9017	26923	116166	575263	3654

Table 2.8: Wage and overtime premium Model (income group sensitivity)

1. **Model:** $\log w_{it} = -\alpha \log(h_{it}) + \beta_0 \frac{h_{it}-\bar{h}}{\bar{h}} * RuleChange_t + \beta_1 \frac{h_{it}-\bar{h}}{\bar{h}} + X_{it}\gamma + \delta_t + \xi_i + \varepsilon_{it}$
2. RuleChange is time dummy for enactment of 2004-FLSA regulations change. (2004 September)
3. 2004 SIPP Panel, from year 2003 (starting July) to 2007
4. Samples exclude those who work for different jobs at the same time, or self employed
5. “Treated" group refers to jobs spells that are exempted before 2004 rule and not exempted after
6. Standard Errors clustered at job spell level
7. Year-Month and state fixed effect included

Dependent Variable	log hours (monthly)	log hours (weekly)	Overtime Dummy	log income (monthly)	log wage (hourly)	wage (hourly)
Treated*rule	0.019**	0.019**	0.023*	0.088***	0.066***	1.054***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.37)
Exempt*rule	-0.001	-0.001	-0.016**	0.006	0.006	1.504***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.27)
Age	0.008***	0.008***	0.010*	0.018***	0.010	-0.194
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.25)
$Age^2/100$	-0.009***	-0.009***	-0.013**	-0.022***	-0.013*	0.311
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.28)
Edu	-0.001	-0.001	-0.007	0.008	0.008	0.348
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.30)
Union	0.006	0.006	0.002	0.033	0.027	0.705
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.56)
Married	0.001	0.001	0.002	0.000	-0.001	0.008
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.09)
r2	0.005	0.005	0.008	0.024	0.025	0.008
N	358077	358077	358077	355842	355842	355842

Table 2.9: Effect of rule change on hours and income (income group sensitivity)

1. **Model:** $Y_{it} = \alpha_0 Treated_i * RuleChange_t + \alpha_1 Exempt_i * RuleChange_t + X_{it}\beta + \delta_t + \xi_i + \varepsilon_{it}$
2. RuleChange is time dummy for enactment of 2004-FLSA regulations change. (2004 September)
3. 2004 SIPP Panel, from year 2003 (starting July) to 2007
4. Samples exclude those who work for different jobs at the same time, or self employed
5. “Treated" group refers to job spells that are exempted before 2004 rule change and not exempted after
6. Standard Errors clustered at job spell level
7. Year-Month and state fixed effect included

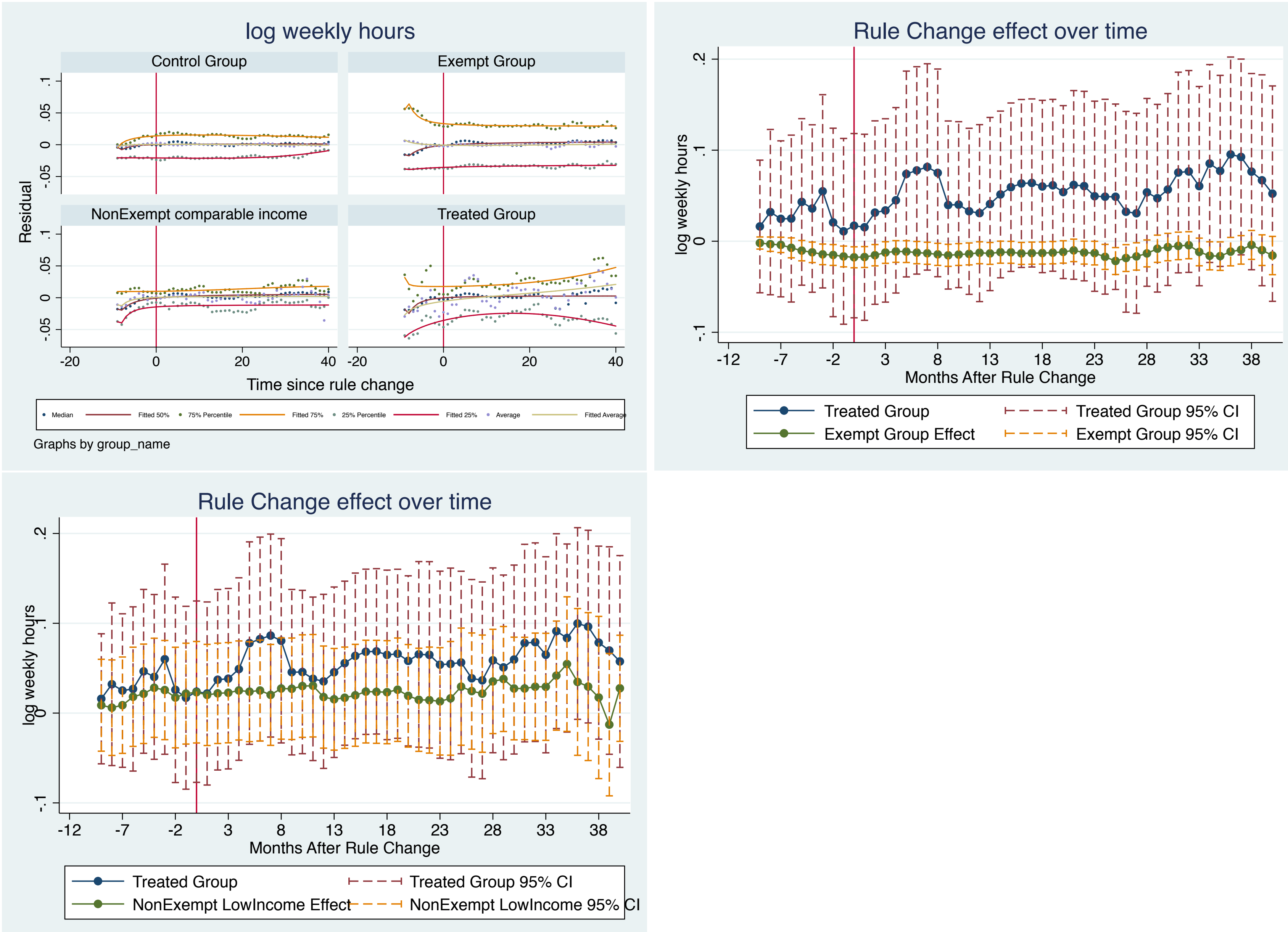


Figure 2.1: Diff-in-Diff Estimates for weekly working hours (SIPP)

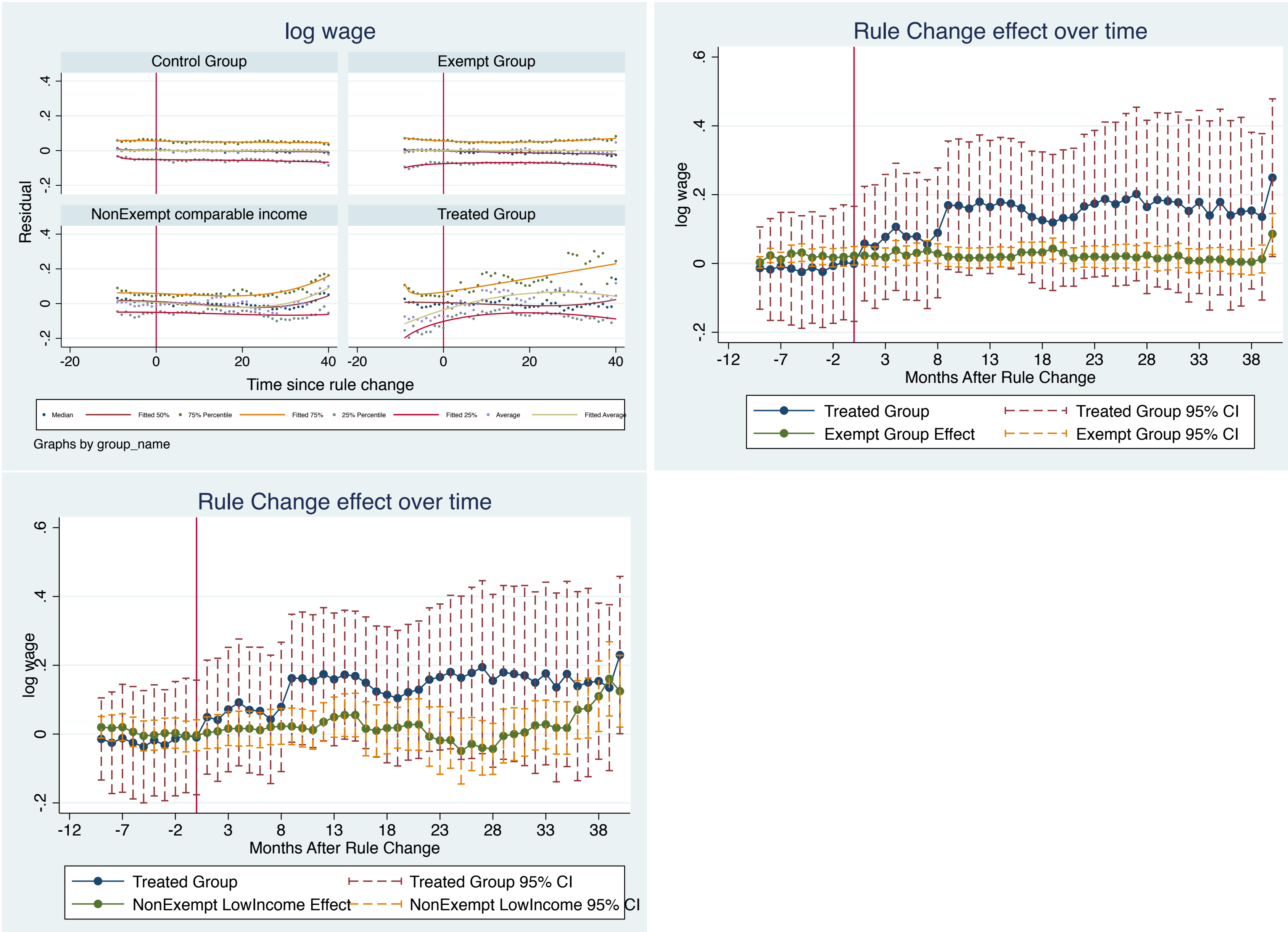


Figure 2.2: Diff-in-Diff Estimates for wage (SIPP)

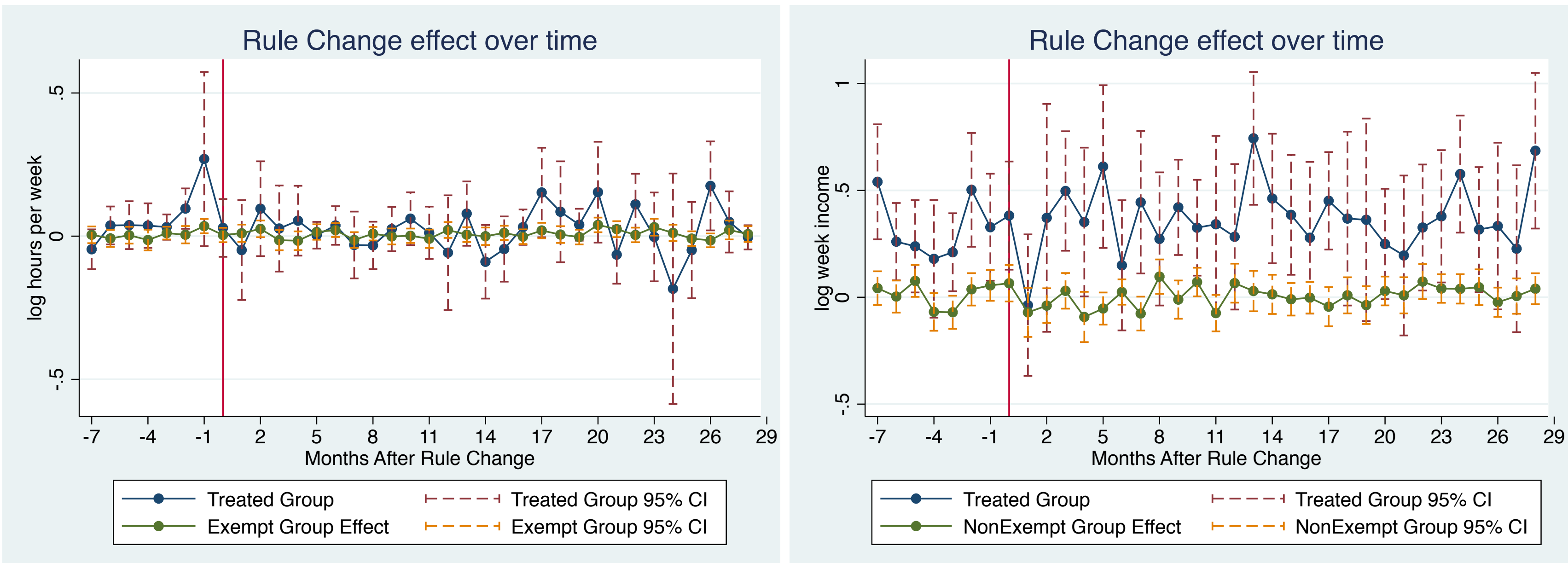


Figure 2.3: Diff-in-Diff Estimates for working hours and wage (CPS)

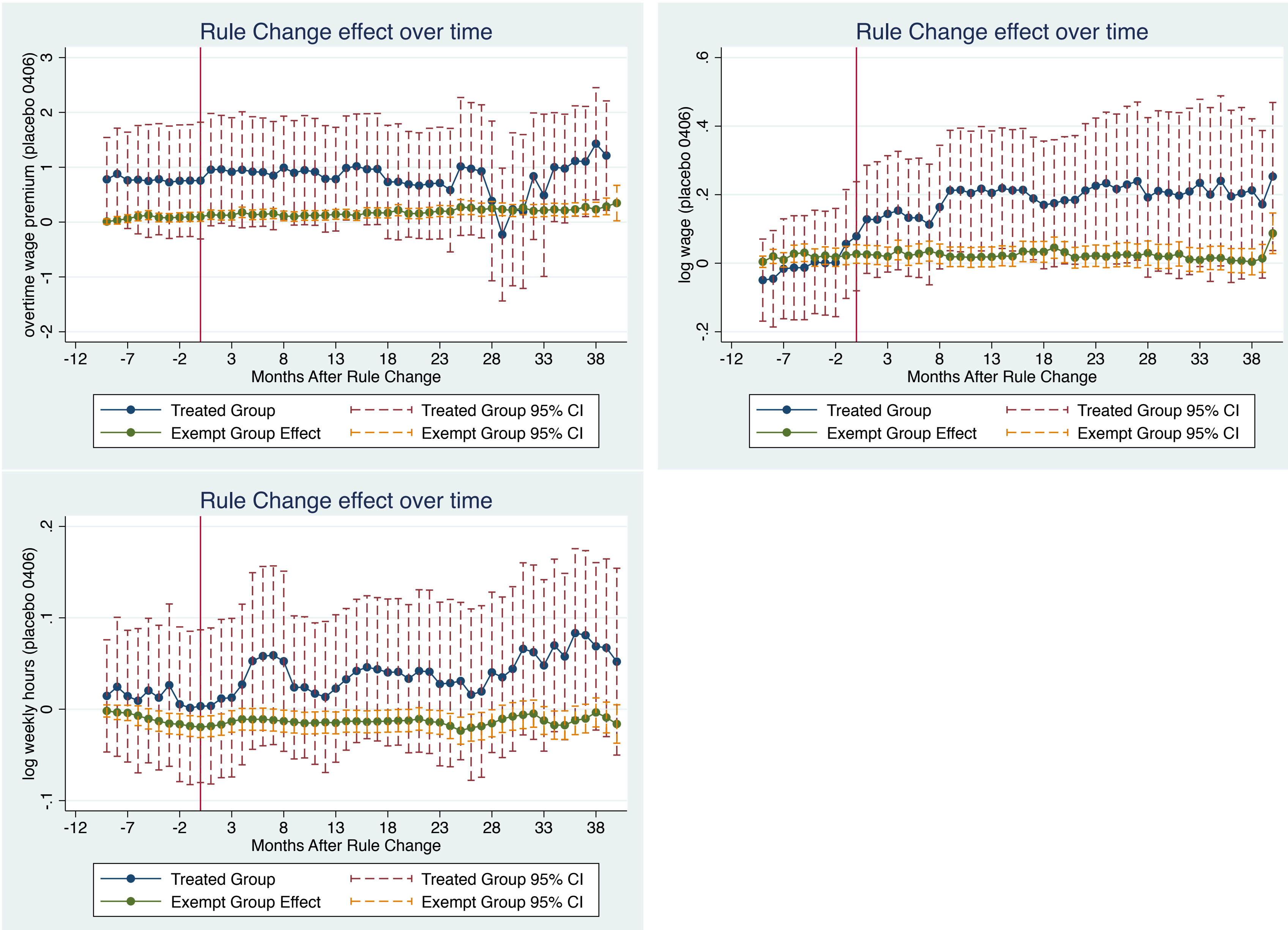


Figure 2.4: Rule change placebo test

Chapter 3

Creditor Governance and Corporate Innovation

3.1 Introduction

The recent Great Recession has again drawn the attention of economists and policy makers alike to the role of financial intermediaries, in particular, banks in the aggregate economy, spurring a wide range of works from macro study on financial stability and macro-economic policy to micro analysis of firm-level credit provision, investment and employment growth. Banks not only provide ex ante screening and credit to firms, but also conduct ex post monitoring to and renegotiations with the borrowing firms. Recently a growing literature in corporate finance demonstrates that banks have an effective monitoring role on firms well before and beyond the bankruptcy state and documents a positive effect of this creditor governance. Their focus is on debt covenants, a non-pricing loan contractual arrangement designed specifically to mitigate the creditor and shareholder conflicts of interests. When the firm violates the covenant, it triggers an effective technical default state and grants the bank legal rights to accelerate or terminate the loan and reduce unused loan commitment. Banks usually renegotiate with the firms and obtain partial control rights with this threat. The recent literature shows firms that violate debt covenants reduce internal and external investment, cut costs and more likely to fire the CEO. It is also found, unsurprisingly, that firm covenant violation is counter cyclical, with more firms violating covenants in economic recessions.

This paper empirically examines the effect of creditor governance, defined as increased creditor (bank) monitoring and control rights in the state of debt covenant violation on corporate innovation. Understanding the role of creditors on innovation is important because innovation, by affecting productivity, is a fundamental driver of economic growth (Solow (1957)) and bank loans have been long documented as the predominant source of external finance around the world, including the US (Gorton and Winton (2003)). The goal of this study is to answer these questions by examining the response of corporate financial policies to covenant violations.

Covenant violations provide a great setting in which economists could examine the interaction of incentive conflicts between debt-holders and shareholders, the creditor governance and corporate innovation for the following number of reasons. First, covenants are a special mechanism designed in contract with the purpose of mitigating the incentive conflicts between managers and creditors, documented in the classic paper by (Smith Jr and Warner (1979)). Second, creditors obtain the right to demand immediate repayment and withholding of further credit in the state of debt covenant violation. This means the creditor has increased bargaining power and could monitor the firm's behavior. A recent burgeoning stream of papers document that the increased creditor right due to covenant violation caused the firm to cut investment, lowered acquisition value, and decreased net debt and stock issuance (Nini, Smith and Sufi (2012)). Also the CEO is more likely to be fired due to covenant violation. Third, covenant violations occur frequently (Dichev and Skinner (2002)) and occurs well before the final stage of firm bankruptcy (Gopalakrishnan and Parkash (1995)), thus making it an important and useful source of variation of firm's financial situation. Lastly, the particular nature of a covenant violation enables one to employ a quasi regression discontinuity design that helps identify the causal effect of violations on corporate innovations.

Theoretically, an increased creditor right could likely have adverse impact on a firm's corporate innovation activities. Conducting innovative activities, unlike routine tasks such as production and sales, requires willingness to take on risk and endure the possibility of failure. Unlike shareholders who enjoy the up-side of the firm, creditors will have to do with the losses if the innovative projects take up a large amount of firm-level resources and end up without commercial value. Managers may also have a risk-shifting incentive to divert projects to risky and innovative activities. In this view, since covenants are designed specifically to deal with the conflict of interest between shareholders and creditors, it would not be surprising that an increased creditor right will make the firm to choose less risky investment strategy and reduce the level of innovative output. On the other hand, the existing literature on the role of creditor governance demonstrates that when the firm violates debt covenants, the increased creditor rights has effective power in disciplining the manager. This view acknowledges the more prominent issue is the agency problem between shareholders and managers instead of the conflict of interests between shareholders and creditors. It therefore posits that the increased monitoring of managers will induce more efficient usage of resources within the firm and could likely increase the overall efficiency of innovative output. It is clearly that how the increased creditor power due to covenant violation turns out to be an empirical question.

Our results show that when a firm violates its debt contract covenants, the firm will reduce investment in capital expenditure and face difficulties in raising new capital either from debt market or stock market, replicating existing literature. However, in terms of firm's R&D spending, there is neither statistically significant nor quantitatively important change due to covenant violation. For our main results on the number of two-year forward patents filed and finally granted as well as the total citation count of these patents, we find there is a statistically significant decrease in the number of patents. The economic magnitude of the effect is modest, at about 3% of the average patents. When we turn to look at the overall citations of the patents, we find a statistically significant 13% increase in the citations to these patents. When we look at the dynamic effects of the covenant violation, we find the results to be quite persistent for up to 3 years after the initial year of covenant violation. The effects gradually increase over the lag year for both the raw patent count and citations count, finally reaching at 7% decrease for the patents and 30% increase for the citations. Results from our robustness checks also confirm the validity of our main results. These results are consistent with the view that an increased creditor power will not necessarily harm firm innovation, but rather, it increases innovation efficiency by cutting managerial slack and induce more efficient allocation of resources inside the firm.

The main contribution of this project will further our understanding on the importance of bank monitoring. Our work joins a burgeoning works that focus on the real effect of debt holder control rights in the setting of covenant violations (e.g., Chava and Roberts (2008); Roberts and Sufi (2009); Nini, Smith and Sufi (2012)). These works in general find an active role played by creditors in corporate governance, and uncover substantial improvement in firm performance associated with this creditor control, for example, relieving investment distortion (Chava and Roberts (2008)), replacing nonperforming managers, and improving subsequent operating and stock price performance (Nini, Smith and Sufi (2012)).

Our paper also belongs to the literature that discusses the financial and institutional determinants of innovation. An emerging empirical evidence shows that institutional aspects and players in financial markets are all important determinants of corporate innovation, for example, stock liquidity (Fang, Tian and Tice (2013)), investment cycles in financial markets (Nanda and Rhodes-Kropf (2013)), financial analysts (He and Tian (2013)), product market competition (Aghion et al. (2005)), investors tolerance for failure (Tian and Wang (2014)), financial development in a local region or country (Hsu, Tian and Xu (2014)), corporate governance and takeover regulation (Atanassov (2013)), banking deregulation (Amore, Schneider and Žaldokas (2013),Chava et al. (2013)), banking competition (Cornaggia, Tian and Wolfe (2012)), and institutional ownership (Aghion, Van Reenen and Zingales (2013)), private rather than public ownership (Ferreira, Manso and Silva (2012)) all affect innovation. We contribute to this literature by showing that increased creditor power in the state of covenant violation are also a first order determinant.

This paper is organized as the following. Section 2 introduces our main dataset and discusses a novel pattern that is the file to grant year lag has been increasing over the years. Section 3 discusses our empirical strategy and spells out our empirical specifications. Section 4 discusses our empirical results including robustness checks. Section 5 concludes.

3.2 Data

3.2.1 Main Dataset

The sample construction consists of all Compustat firm from 1996 to 2005, excluding fnancial firms (SIC codes 6000-6999) and utility firms (SIC codes 4900-4999). The beginning year of the sample coincides with information provided by Nini, Smith and Sufi (2012) on covenant violations, which only starts from 1996 because electronic filing first became available for all SEC-registered firms. Nini, Smith and Sufi (2012) use a textual-search based analysis to identify the existence of covenant violations from 10-K and 10-Q Securities and Exchange Commission (SEC) filings. Their database is available for public usage on Amir Sufi's website[1]. Interested readers could consult the Data Appendix of Nini, Smith and Sufi (2012) for more details on the search algorithm and data information. To construct the sample, I merge this covenant-violation dataset with Compustat firm observations.

The firm-level data of patents and citations are provided by Kogan et al. (2012)[2]. This database contains utility patents issued by the United States Patent and Trademark Office (USPTO) between January 1, 1926 and November 2, 2010, along with citation data on those patents. Following Nini, Smith and Sufi (2012), I use a set of variables called "covenant controls" to account for the accounting ratios on which financial covenants are written and at the same time proxy for firm-specific financial condition and investment opportunities and thus could very well correlate with my

[1] http://faculty.chicagobooth.edu/amir.sufi/data.html
[2] https://iu.box.com/patents

outcome of interest. These variables include: lagged book debt to assets ratio, lagged cash to assets ratio, lagged EBITDA to assets ratio, lagged cashflow to assets ratio and current interest expense to lagged assets ratio. Other controls include logarithm of book assets, tangibility, cash holding, ROA, net capital expenditure, R&D spending and Tobin's Q. All variables used in this study are formally defined in Table 4. I require all these variables to be non-missing and thus end up with 58883 firm-year observations. To mitigate the impact of outliers on my analysis, I winsorize all variables at the 5th and 95th percentiles, and my results remain quantitatively similar if I winsorize at the 1st and 99th percentiles.

Table 5 lists the summary statistics for variables. The outcome variable, patent count for each firm in each year, has an average of 11 and a median of 0, suggesting that the distribution of patent created by firms are highly skewed. The mean and median of citation count for each firm and year is 97 and 0, respectively. It is worth noting that the patent counts are quite similar but the patent citation numbers are much larger than other papers that mainly rely on the NBER Patent Citation Database in concurrent period. For example, He and Tian (2013) has an average patent count of only 4 citations per patent on average, while that number more than doubles to 9 citations per patent. This is because the Kogan et al. (2012) dataset use google patent database and are able to recover more patents and their citations.

3.2.2 The increasing file-grant lag

An interesting by-product from this project is the finding that in the US, the number of years between the year the patent is filed or applied and the year the patent is granted (the file-grant lag) has been steadily increasing over the past 10 years, a finding not recorded in the existing literature on innovation. To arrive at this finding, I dropped those patents whose grant year is earlier than the file year as well as those patents whose grant year are missing. I also calculate the file-grant lag as the difference of the grant year and the file year, with the caveat that some patents may be filed at the end of a calendar year and granted in the beginning of a calendar year. In Table 1, I present the proportion of patents finally granted by the year-bin (say, 1996 and 1997) of the filing years with respect to the number of file-grant lag from 0 to 7 and more years. From 1996 to 1999, the proportion of patents with 4 lag year is still not over 10% and the median lag year is 2. This pattern is consistent with the report by Hall, Jaffe and Trajtenberg (2001) in which they tabulate the filed and finally granted patents in years from 1967 to 1992 and find a somewhat consistent pattern a median or mean lag year between 2 and 3 years. However, once we look beyond 1999, the distribution of the file-grant lag changes substantially. The proportion of patents with lag-year of 4 years increased continuously from 11% in 2000-1 year-bin to 21% in 2004-5 year-bin. The median lag-year also increases from 2 to 3 years. Finally in the year-bin of 2004-5, there is a substantial fraction of patents, in fact 16%, are granted 5 years after the filing date. I also tabulate the lag-year distribution for years after 2005, but since the dataset ends in 2010, many of the patents filed in these years are still truncated, making further inference not feasible.

Table 2 tabulates the distribution of lag-years according to the patent grant years. This different way of tabulation from Hall, Jaffe and Trajtenberg (2001) has the benefit of providing information on the newest observations. We see that, for example, there is a non-negligible fraction of patents granted in 2010 that were filed 5 or more than 5 years ago, the fraction being over 30%. Rolling back to 1996-7 year bin, the number of patents that were filed 5 or more than 5 years account for less than 4%. I also tabulate the mean and standard deviation of the lag-year of patents by either the file year or the grant year in Table 3. Both the mean and standard deviation are almost monotonically increasing over time, with the stronger case made by examining the grant-year.

Evidently, a detailed investigation of why there is such a changing distribution of the file-grant lag in the patent warrants a project by itself and is beyond the scope of this paper. It is worth noting the implications of this changing empirical distribution. The truncation problem for the NBER patent database has been well documented in Hall, Jaffe and Trajtenberg (2001). Usually papers argue that the file-grant lag is only 2 or 3 years and therefore one approach is to simply drop the last two or three years' observations in the patent database used. The implication is that one should drop the last 5 years' observations, if one wants to use some recent patent databases. An alternative to correct for the truncation bias in patent counts is to use the weight factors computed from the application-grant empirical distribution, advocated by Hall, Jaffe and Trajtenberg (2001). This methodology works when the empirical distribution of the file-grant lag is stable, and clearly with recent data, the empirical distribution is changing and may make using this approach unsuitable.

3.3 Empirical Strategy

3.3.1 The Covenant Violation

The main empirical concern for an investigation of the effects of covenant violation is omitted variable bias. That is, a firm violating covenant must not be performing well, compared with a firm that does not, therefore the violating firm lacks the investment opportunities to do more innovation. A naive strategy of comparing the innovative performance of firms violating covenants with those that do not will not work. To disentangle the effects of the covenant violation from changes in corporate innovation that would have otherwise occurred, I resort to the use quasi-discontinuity regressions following Roberts and Sufi (2009) and Nini, Smith and Sufi (2012), which exploit the discontinuity created at the point of violation. To be more precise, covenant violations depend exclusively on whether a known variable is below or above a pre-specified threshold, regardless of the distance to the threshold. This rule arbitrarily creates a discontinuous treatment in the neighborhood of a known cut-off, and thus resembles a randomized trial. We apply this discontinuity approach by including as right-hand side variables a covenant violation indicator variable along with linear and nonlinear functions of the underlying variables on which covenants are written. With this empirical specification, the point estimate on the covenant violation indicator variable is identified under the assumption that corporate innovation decisions and performance are not discontinuous exactly at the covenant threshold. As long as firm's innovation performance (or unobserved variables that affect innovation such as investment opportunities), as a function of variables on which covenants are written, does not exhibit the same exact discontinuity at the covenant threshold, we are able to identify the effect of covenant violation as a discrete change on corporate innovation.

3.3.2 Empirical Specifications

Our right hand side variables used in our empirical specification are composed of financial and accounting variables on which covenants are written. Following Nini, Smith and Sufi (2012), the choice of these controls is based on the most common financial covenants employed in debt contract agreements. These variables include: lagged book debt to assets ratio, lagged cash to assets ratio, lagged EBITDA to assets ratio, lagged cashflow to assets ratio and current interest expense to lagged assets ratio. Other controls include logarithm of book assets, tangibility, cash holding, ROA, net capital expenditure, R&D spending and Tobin's Q, following the literature on corporate innovation, such as Aghion, Van Reenen and Zingales (2013). We use up to 4th order of these variables to flexibly control for the nonlinear effects these variables could have on our outcome of interests.

A large sample of violating and non-violating firms are estimated in a dynamic model of corporate innovation, with right-hand side variables including a covenant violation indicator variable along with linear and nonlinear functions of the covenant controls mentioned above. By flexibly controlling for these variables, the impact of a covenant violation is identified by the discontinuity occurring at the covenant threshold. More specifically, we estimate the regressions as follows:

$$Y_{i,t+2} = \alpha + \beta_1 Violation_{i,t} + X_{i,t}\beta_2 + Z_{i,t}\beta_3 + Z_{i,t-1}\beta_4 + HigherOrders_{i,t}\beta_4 + firm_i + year_t + \varepsilon_{i,t}$$

where our outcome variable $Y_{i,t+2}$ is either the patent count or the total citations count. Due to the nature of innovation and patenting, we use the outcome variable at two years later. Violation is an indicator variable that equals 1 for a financial covenant violation, X is a vector of innovation control variables, Z is a vector of covenant control variables, we use the current and lagged covenant control variables and HigherOrders is the up to 4th order series of the each of the covenant violation controls. Lastly, we include firm fixed effects to control for any firm-invariant properties in innovation and year fixed effects to account for aggregate time trends.

3.4 Empirical Results

3.4.1 Results on corporate real and financial policies

Before discussing the main results of the covenant violation on firm innovation activities, we first examine the impact on the firms' real and financial activities in terms of capital expenditure and R&D spending, net debt issuance and equity issuance.Existing literature has shown that following covenant violation, firms sharply cut their capital expenditure and face great difficulty in obtaining new loans as well as issuing new stocks (See Nini, Smith and Sufi (2012)). The main purpose of our empirical exercise is to demonstrate the validity of my empirical approach by replicating the existing papers. At the same time, we also contribute a novel evidence to the literature by examining the firm's R&D spending.

Table 6 shows the effect of covenant violation on capital expenditure (investment), R&D spending, net debt issuance and stock issuance. In the odd columns, our specifications only include the violation indicator and the firm and year fixed effects. We add the full controls in the even columns. Our results show that there is a significant drop in the capital expenditure, with a point estimate of -0.026 and is significant at 1% level. In terms of economic magnitude, compared with an average investment rate of 0.36 it amounts to a 7.6% drop in investment. Similarly, we also find significant drop in net debt issuance and stock issuance. When we examine the effect on R&D spending, we find there is no effect due to covenant violation in terms of the amount of spending, with a point estimate of almost identical to 0 and not significant at any reasonable level of confidence. This finding demonstrates that even though firms reduce investment in capital, face difficulties in raising new capital either through net debt issuance or new stock issuance, there is no change in firms' spending on research and development. However, we think of this as some preliminary evidence on corporate innovation as the accounting rules allows flexible discretion in allocating resources used in actual R&D activities.

3.4.2 Main Results on Innovation

In this section, we examine the impact of covenant violation on innovation in terms of the number of patents filed and finally granted and the total citation count of these patents 2 years forward. We use the 2-year forward outcome variable because of the nature of technological innovation is slow-moving. Our results are robust to examining either 1-year forward or 3-year forward outcome variables, with similar quantitative results in these specifications. We present the results of our quasi- regression discontinuity design in Table 7. When we regress the number of patent count on the covenant violation dummy and firm and year fixed effects, we already see a statistically significant drop in the number of patents. In the full specification where we fully control for the confounding effects surrounding covenant violations, we find the point estimate to be -0.4 and significant at 1% level of confidence. Given an average patent number of 11 this account for about 3% drop of the total patents. This modest decline in the number of patents suggest that an increased creditor power will not have substantial adverse effects on corporate innovation. When we examine the total counts of the number of citations, we find quite surprisingly, there is a positive and statistically significant increase in the total number of citations, with the point estimate being 13.5 and significant at 1% level of confidence. This means an economic magnitude of about 14% increases in total citations, given an average citation number of 98. Results imply that the quality of patents has in fact increased, which suggests a more efficient use of organizational resources to promote innovation, despite the fact that the firm has cut spending in capital expenditure and faces difficulties in raising capital from creditors or shareholders. Contrary to one view in the literature that creditors are anti-innovation due to the riskiness of firm R&D and the potential risk shifting behavior by the managers on behalf of shareholders, we actually find in the state of covenant violation, the firm has not substantially reduced patent number but increased the quality of patents. Results imply that increased creditor monitoring could indeed discipline managers to make the most efficient use of resources to conduct innovative activities.

3.4.3 Dynamic Effects of Covenant Violation

In this section, we proceed to examine the dynamic effects of the covenant violation, trying to investigate if such a causal effect we identified in the earlier subsection is persistent or short-lived. To do so we add in our main specification the 1-year, 2-year and 3-year lagged covenant violation indicator. For example, the 3-year lagged indicator identifies the effect covenant violation has on the outcome of interest 3 years later, for the 2-year forward innovation outcome. If the effects are short-lived, we would expect these lag-year effects to be decreasing in absolute magnitude. Results for these empirical exercises are presented in Table 8. We find the effects of covenant violation for the total patent number increased slightly from -0.42 (the immediate effect) to -0.64 (3-year after covenant violation), although the coefficients are not statistically significant from each other. We also examine the total citation counts, and we find a relative larger increase in the effects three years after covenant violation. The magnitude doubled from 11 to 27 which could account for 30% of the average total citations.

3.4.4 Robustness

Finally we conduct a batch of robustness checks to confirm the validity of our main results. The results are presented in Table 9. First, in column 1 we cluster the standard error at industry level rather than at firm level as in the main regressions. The results are still significant at 5% level with the same point estimate but an elevated standard errors. We also use bootstrap to generate standard errors to account for the serial correlations in standard errors and as shown in column 2, the results are still significant at 1% level of confidence. We use the 1-year forward patent count and total citations count in column 3 and 4 and find the results are still statistically significant and quantitatively similar to our main results. Since we are dealing with count data for both our patent count and citation count, we use fixed effects poisson model and negative binomial model separately in column 5 and 6 and results are again statistically significant at 1% level of confidence.

3.5 Conclusion

In this paper, we exploit the institutional feature that in the state of debt contract covenant violation, creditors obtain increased control right of the firm to examine the effect of increased creditor governance well before the state of bankruptcy on corporate innovation input and outcome. Theoretical predictions are ambiguous, as the conflict of interests between creditors and shareholders could justify an increased creditor right leading to lower innovative performance while at the same time creditor monitoring could discipline managers and increase the efficiency of corporate innovation. Consistent with the view that increased creditor monitoring has disciplining effect on the managers, we find no significant change in the R&D spending, significant but model decrease in the total patent counts two years forward as well as significant and large positive impact on the citation counts of the patents. The results we identified are persistent up to three years after the initial year of covenant violation. Our results demonstrate that increased creditor governance is overall benefitial to firm innovation. Future works can be done to explore the specific mechanisms leading up to the change in the internal organization and resource reallocation of the firm.

Table 1: The changing distribution of file-grant lag in filing year

This table describes the changing distribution of the lag between filing year and grant year using the Google Patent Database collected by Kogan et al. (2012) and available at https://iu.box.com/patents . I dropped those patents whose grant year is earlier than the file year and whose grant year is missing. Each column tabulate the percentage of patents filed in that year and finally granted with respect to their file-grant year lag from 0 to 7 more years. There is a clear pattern of increasing lag years between filing and grant. The lag years for recent filing patents are missing as the database ends in 2010.

See Hall et al. for a tabulation of patents applied in years from 1967 to 1992 for comparison, where there is no such a pattern of increasing lag years.

				Patent Filing Years					
	1996-7	1998-9	2000-1	2002-3	2004-5	2006-7	2008-9	2010	Total
lag years				**Distribution of Lags in %**					
0	0.53	0.59	0.63	0.69	0.4	0.79	2.11	100	0.65
1	20.73	19.6	17.88	14.47	9.2	12.03	37.68	0	16.37
2	47.8	41.09	34.56	28.57	24.37	28.61	60.21	0	34.82
3	20.99	22.11	19.87	21.39	24.85	39.22	0	0	23.11
4	5.98	9.03	11.11	15.61	20.81	19.34	0	0	13.05
5	2.07	3.71	7.95	9.58	15.58	0	0	0	6.87
6	0.97	1.74	4.22	5.59	4.79	0	0	0	3.11
7+	0.93	2.12	3.79	4.1	0	0	0	0	2.02
Total	100	100	100	100	100	100	100	100	100

Table 2: The changing distribution of file-grant lag in grant years.

This table describes the changing distribution of the lag between filing year and grant year using the Google Patent Database collected by Kogan et al. (2012) and available at https://iu.box.com/patents . I dropped those patents whose grant year is earlier than the file year and whose grant year is missing. Each column tabulate the percentage of patents granted in that year with respect to the lag years from 0 to 7 more years. Again, there is a clear pattern that the file-grant lags have been increasing over the years.

	Patent Grant Years								
	1996-7	1998-9	2000-1	2002-3	2004-5	2006-7	2008-9	2010	Total
Lag (years)	**Distribution of Lags in %**								
0	0.75	0.69	0.78	0.81	0.44	0.53	0.32	0.21	0.58
1	30.12	23.08	21.7	20.6	13.08	9.26	7.18	4.38	16.21
2	51.18	49.91	44.32	41.61	36.73	25.19	19.54	16.23	35.89
3	14.16	19.96	23.67	23.12	25.49	27.07	25.76	24.14	23.26
4	2.47	4.41	5.81	9.37	14.6	18.67	21.91	23.07	12.43
5	0.75	1.02	1.83	2.61	6.6	11.33	13.54	15.78	6.46
6	0.29	0.52	0.95	0.91	1.97	5.17	6.81	8.8	3
7+	0.28	0.4	0.94	0.96	1.1	2.77	4.93	7.39	2.16
Total	100	100	100	100	100	100	100	100	100

Table 3: **Mean and Standard Deviation of the file-grant lag for each filing and grant year.** This table describes the mean and standard deviation of the lag between filing year and grant year using the Google Patent Database collected by Kogan et al. (2012) and available at https://iu.box.com/patents . I dropped those patents whose grant year is earlier than the file year and whose grant year is missing. The left panel is for the filing year and the right panel is for the grant year. The increase in the file-grant lag is almost monotonic when organized in the grant year.

File Year	Mean	Std. Dev.	Freq.	**Grant Year**	Mean	Std. Dev.	Freq.
1996	2.246049	1.117476	144886	1996	1.854944	1.112238	109654
1997	2.28861	1.166078	169582	1997	2.008338	0.954982	112020
1998	2.429695	1.283346	168473	1998	2.092572	0.964524	147572
1999	2.564102	1.475243	180673	1999	2.147105	1.044452	153591
2000	2.801871	1.643921	196528	2000	2.236135	1.173776	157596
2001	2.904018	1.692387	208665	2001	2.307521	1.251668	166158
2002	2.994196	1.659626	205735	2002	2.300908	1.228359	167400
2003	3.197978	1.614457	190536	2003	2.439945	1.251298	169078
2004	3.288417	1.444881	179785	2004	2.581772	1.305458	164384
2005	3.15235	1.264196	162593	2005	2.943673	1.385309	143891
2006	2.842217	1.019163	132917	2006	3.184933	1.521158	173825
2007	2.346354	0.746353	89244	2007	3.307895	1.549033	157336
2008	1.714792	0.479713	41517	2008	3.479607	1.609778	157796
2009	0.94098	0.235676	8675	2009	3.663876	1.657468	167468
2010	0	0	397	2010	3.907035	1.732287	186243
Total	2.758561	1.459228	2080206	Total	2.74756	1.501151	2334012

Table 4: Variable Definition

Measures of Innovation	
Patent_t+2	The total number of patents filed and eventually granted of company in two years ahead.
Cite_t+2	The total number of non-self-citations received on the firm's patents filed and eventually granted of company in two years ahead.
Measures of Firm Variables	
Assets_t	Log of book value of total assets (Compustat data item ♯6) measured at the end of fiscal year t
R&DAssets_t	Research and development (R&D) expenditure (♯46) divided by book value of total assets (♯6) measured at the end of fiscal year t, set to 0 if missing
ROA_t	Return on assets ratio defined as operating income before depreciation (♯13) divided by book value of total assets (♯6), measured at the end of fiscal year t.
Investment_t	Property, plant & equipment (♯8) divided by book value of total assets (♯6) measured at the end of fiscal year t
Leverage_t	Firm i's leverage ratio, defined as book value of debt (♯9+♯34) divided by book value of total assets (♯6) measured at the end of fiscal year t
CapexAssets_t	Capital expenditure (♯128) scaled by book value of total assets (♯6) measured at the end of fiscal year t
TobinQ_t	Firm i's market-to-book ratio during fiscal year t, calculated as market value of equity (♯199♯25) plus book value of assets (♯6) minus book value of equity (♯60) minus balance sheet deferred taxes (♯74, set to zero if missing), divided by book value of assets (♯6)
RDD Controls	
bkDebtAssets_t	The lagged book debt to assets ratio, defined as the lagged value of book value of debt (♯9+♯34) divided by current value of book value of total assets (♯6).
CashAssets_t	The lagged cash to assets ratio, defined as the lagged value of cash (♯308) divided by current value of book value of total assets (♯6).
ebitdaAssets_t	The lagged EBITDA (♯18) divided by the current value of book value of total assets (♯6).
CashFlowAssets_t	The lagged cashflow ((♯8+♯5) divided by the current value of book value of total assets (♯6).
InterestAssets_t	The current interest expense (♯15) to lagged value of book value of total assets (♯6).

Table 5 Summary Statistics for All Variables Used in the Paper

S.D. is sample standard deviation of the corresponding variable while p10(/50/90) is the 10(/50/90)% percentile of the variable of interest. The number of observations for each variable varies based on data availability. Definitions of all variables are in Table 4.

	Obs.	Mean	S.D.	p10	p50	p90
Patent	51182	11.256	105.133	0.000	0.000	7.000
Citation	51182	97.622	1181.285	0.000	0.000	48.000
Asset	71360	4.726	2.434	1.781	4.658	7.871
R&DAssets	71360	0.094	2.506	0.000	0.000	0.200
ROA	71332	-0.026	0.224	-0.461	0.048	0.171
PPEAssets	71347	0.279	0.237	0.039	0.205	0.658
Leverage	71360	0.262	0.323	0.000	0.189	0.574
Investment	70491	0.364	0.399	0.047	0.224	0.880
TobinQ	71360	2.138	1.596	0.864	1.527	4.700
Tangibility	71360	284.544	2651.574	-0.000	1.259	271.623
bkDebtAssets	59602	0.628	27.541	0.000	0.167	0.579
CashAssets	59602	-0.574	27.227	-0.143	0.000	0.025
ebitdaAssets	59602	-0.432	20.600	-0.459	0.022	0.142
CashFlowAssets	59602	-0.752	35.891	-0.453	0.057	0.152
InterestAssets	66027	0.071	1.929	0.000	0.015	0.061

Table 6: Effects of Covenant Violation on capital expenditure (investment), R&D spending, Net Debt Issuance and Stock Issuance.

This table reports results from our Quasi Regression Discontinuity Design estimations for the effect of covenant violation on capital expenditure (investment), R&D spending, Net Debt Issuance and Stock Issuance. Definitions of all other control variables are provided in Table 4. Robust standard errors clustered at firm-level are reported in parentheses. Our sample ranges from 1996 to 2005. Regressions in all columns include year-fixed effects and firm-fixed effects. We include the higher order controls for the covenant violations up to 4-th orders in the even columns. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively.

	(1) investment	(2) investment	(3) R&D spending	(4) R&D spending	(5) Net Debt	(6) Net Debt	(7) Stock Issuance	(8) Stock Issuance
Violation	-0.042***	-0.026***	0.005**	0.000	-0.023***	-0.035***	-2.239***	-1.255***
	(0.005)	(0.004)	(0.002)	(0.003)	(0.007)	(0.008)	(0.327)	(0.371)
Asset		0.092***		-0.052***		0.117***		14.654***
		(0.004)		(0.010)		(0.006)		(2.083)
PPEAssets		0.050**		0.309***		-0.091***		-21.990***
		(0.024)		(0.094)		(0.030)		(6.747)
ROA		0.273***		-0.082***		-0.301***		-13.725***
		(0.018)		(0.025)		(0.023)		(4.416)
Tangibility		-0.000		0.000***		-0.000		0.011
		(0.000)		(0.000)		(0.000)		(0.007)
TobinQ		0.047***		0.006**		0.014***		5.675***
		(0.002)		(0.003)		(0.003)		(0.725)
Higher Order Controls	No	Yes	No	Yes	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	70491	58883	71360	59569	67009	55791	70201	58580
R^2	0.042	0.092	0.000	0.249	0.007	0.027	0.003	0.028

Table 7: Effects of Covenant Violation on the number of patents and total citations.

This table reports results from our Quasi Regression Discontinuity Design estimations for the effect of covenant violation on patent filed (and finally granted) as well as the total number of citations up to 2010 for those patents. Definitions of all other control variables are provided in Table 4. Robust standard errors clustered at firm-level are reported in parentheses. Our sample ranges from 1996 to 2005. Regressions in all columns include year-fixed effects and firm-fixed effects. We include the higher order controls for the covenant violations up to 4-th orders in the even columns. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively.

	(1) Patent	(2) Patent	(3) Citation	(4) Citation
Violation	-0.667***	-0.400***	14.329***	13.481***
	(0.155)	(0.140)	(3.597)	(3.452)
Asset		1.594***		-1.800
		(0.596)		(11.301)
R&DAssets		0.224		-0.817
		(0.164)		(3.829)
PPEAssets		2.118		66.976
		(1.967)		(50.190)
ROA		-2.230		112.731**
		(1.795)		(47.234)
Leverage		-2.003**		28.226
		(0.823)		(17.267)
Tangibility		0.000		-0.028
		(0.000)		(0.018)
TobinQ		1.069***		-5.373
		(0.262)		(5.581)
CapexAssets		2.503		100.218
		(1.969)		(61.680)
Higher Order Controls	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Firm Fixed Effect	Yes	Yes	Yes	Yes
Observations	51182	41554	51182	41554
R^2	0.004	0.005	0.009	0.014

Table 8: Dynamic Effects of Covenant Violation on Patent number and citations.

This table reports results from our Quasi Regression Discontinuity Design estimations for the dynamic effect of covenant violation (up to 3 years later) on patent filed (and finally granted) as well as the total number of citations up to 2010 for those patents. Definitions of all other control variables are provided in Table 4. Robust standard errors clustered at firm-level are reported in parentheses. Our sample ranges from 1996 to 2005. Regressions in all columns include year-fixed effects and firm-fixed effects. We include the higher order controls for the covenant violations up to 4-th orders in the even columns. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively.

	(1) Patent	(2) Patent	(3) Citation	(4) Citation
Violation	-0.392**	-0.426**	10.139***	11.364***
	(0.156)	(0.197)	(3.488)	(3.476)
Violation_lag1	-0.498***	-0.407**	13.287***	13.927***
	(0.154)	(0.162)	(4.353)	(4.416)
Violation_lag2	-0.725***	-0.569***	12.784***	13.240***
	(0.183)	(0.185)	(4.520)	(4.825)
Violation_lag3	-0.859***	-0.648**	26.805***	27.154***
	(0.269)	(0.273)	(7.107)	(7.463)
Asset		2.414***		28.194*
		(0.761)		(15.388)
R&DAssets		1.128***		6.845
		(0.345)		(7.948)
PPEAssets		0.239		140.868
		(3.151)		(87.723)
ROA		-4.225*		48.919*
		(2.262)		(27.728)
Leverage		-1.738*		3.466
		(0.912)		(23.008)
Tangibility		0.000		-0.027
		(0.000)		(0.018)
Investment		-0.531		-12.042
		(0.481)		(9.319)
TobinQ		0.385		14.200***
		(0.271)		(5.233)
CapexAssets		0.005		80.747
		(4.251)		(56.706)
Higher Order Controls	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Firm Fixed Effect	Yes	Yes	Yes	Yes
Observations	26693	26372	26693	26372
R^2	0.005	0.006	0.008	0.014

Table 9: Robustness Tests of Covenant Violation on Patent number and citations.

This table reports robustness tests for our Quasi Regression Discontinuity Design estimations for the effect of covenant violation on patent filed (and finally granted) as well as the total number of citations up to 2010 for those patents. The first column we cluster standard error at industry level and we use bootstrap to calculate the standard error in the 2nd column. We use the 1-year later patent and citations in column 3 and 4. We use fixed effects poisson model and negative binomial model in column 5 and 6. Definitions of all other control variables are provided in Table 4. Our sample ranges from 1996 to 2005. We include the higher order controls for the covenant violations up to 4-th orders in the even columns. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively.

	(1) Patent	(2) Patent	(3) npat1	(4) cite_num1	(5) Patent	(6) Patent
Violation	-0.400**	-0.400***	-0.503***	8.863***	-0.145***	-0.100***
	(0.195)	(0.123)	(0.145)	(2.387)	(0.019)	(0.034)
Asset	1.594**	1.594***	1.970***	13.765	0.299***	0.116***
	(0.730)	(0.617)	(0.557)	(9.290)	(0.005)	(0.008)
R&DAssets	0.224	0.224	0.283*	0.180	0.357***	0.016
	(0.144)	(0.191)	(0.146)	(2.881)	(0.026)	(0.054)
PPEAssets	2.118	2.118	2.486	80.273	0.634***	-0.303***
	(2.046)	(2.293)	(2.099)	(52.547)	(0.038)	(0.098)
ROA	-2.230	-2.230	-4.454**	36.393	0.023	0.063
	(2.692)	(1.778)	(2.126)	(25.639)	(0.024)	(0.066)
Leverage	-2.003**	-2.003**	-1.638**	26.324	-0.465***	-0.245***
	(0.861)	(0.835)	(0.786)	(17.275)	(0.020)	(0.054)
Tangibility	0.000	0.000	0.000	-0.030*	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.017)	(0.000)	(0.000)
TobinQ	1.069***	1.069***	0.886***	3.596	0.032***	0.047***
	(0.315)	(0.217)	(0.219)	(3.345)	(0.002)	(0.006)
CapexAssets	2.503	2.503	0.478	33.101	1.359***	0.600***
	(1.718)	(2.512)	(1.049)	(23.160)	(0.059)	(0.200)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	41554	41554	49616	49616	22437	22437
R^2	0.005	0.005	0.005	0.015		

Bibliography

Abowd, John M. and David Card. 1989. "On the Covariance Structure of Earnings and Hours Changes." *Econometrica* 57(2):411–445.

Acemoglu, Daron and David Autor. 2010. "Skills, Tasks and Technologies: Implications for Employment and Earnings." 4.

Aghion, Philippe, John Van Reenen and Luigi Zingales. 2013. "Innovation and Institutional Ownership." *American Economic Review* 103(1):277–304.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith and Peter Howitt. 2005. "Competition and Innovation: an Inverted-U Relationship." *The Quarterly Journal of Economics* 120(2):701–728.

Altonji, Joseph G., Anthony A. Smith and Ivan Vidangos. 2013. "Modeling Earnings Dynamics." *Econometrica* 81(4):1395–1454.
URL: *http://dx.doi.org/10.3982/ECTA8415*

Amore, Mario Daniele, Cédric Schneider and Alminas Žaldokas. 2013. "Credit supply and corporate innovation." *Journal of Financial Economics* 109(3):835–855.

Atanassov, Julian. 2013. "Do hostile takeovers stifle innovation? Evidence from antitakeover legislation and corporate patenting." *The Journal of Finance* 68(3):1097–1131.

Atkinson, A.B. and T. Piketty. 2007. *Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*. OUP Oxford.
URL: *https://books.google.com/books?id=ACbQdj0XLqAC*

Autor, David H. and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103(5):1553–97.

Azariadis, Costas. 1975. "Implicit Contracts and Underemployment Equilibria." *Journal of Political Economy* 83(6):1183–1202.
URL: *http://www.jstor.org/stable/1830855*

Bagger, Jesper, Francois Fontaine, Fabien Postel-Vinay and Jean-Marc Robin. 2014. "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics." *American Economic Review* 104(6):1551–96.
URL: *http://www.aeaweb.org/articles?id=10.1257/aer.104.6.1551*

Beaudry, Paul and John DiNardo. 1991. "The Effect of Implicit Contracts on the Movement of Wages Over the Business Cycle: Evidence from Micro Data." *Journal of Political Economy* 99(4):665–688.
URL: *http://www.jstor.org/stable/2937776*

Besley, Timothy and Robin Burgess. 2004. "Can Labor Regulation Hinder Economic Performance? Evidence from India." *The Quarterly Journal of Economics* 119(1):91–134.
URL: *http://www.jstor.org/stable/25098678*

Botero, Juan C., Simeon Djankov, Rafael La Porta, Florencio Lopez-de Silanes and Andrei Shleifer. 2004. "The Regulation of Labor*." *The Quarterly Journal of Economics* 119(4):1339–1382.
URL: *+ http://dx.doi.org/10.1162/0033553042476215*

Bruins, Marianne, James A. Duffy, Michael P. Keane and Jr Anthony A. Smith. 2015. Generalized Indirect Inference for Discrete Choice Models. Economics Papers 2015-W08 Economics Group, Nuffield College, University of Oxford.
URL: *https://ideas.repec.org/p/nuf/econwp/1508.html*

Chava, Sudheer, Alexander Oettl, Ajay Subramanian and Krishnamurthy V Subramanian. 2013. "Banking deregulation and innovation." *Journal of Financial Economics* 109(3):759–774.

Chava, Sudheer and Michael R Roberts. 2008. "How does financing impact investment? The role of debt covenants." *The Journal of Finance* 63(5):2085–2121.

Cornaggia, Jess, Xuan Tian and Brian Wolfe. 2012. "Does Banking Competition Affect Corporate Innovation?" *Available at SSRN 2017928* .

Costa, Dora L. 2000. "Hours of Work and the Fair Labor Standards Act: A Study of Retail and Wholesale Trade, 1938-1950." *Industrial and Labor Relations Review* 53(4):648–664.
URL: *http://www.jstor.org/stable/2696141*

Dichev, Ilia D and Douglas J Skinner. 2002. "Large–sample evidence on the debt covenant hypothesis." *Journal of Accounting Research* 40(4):1091–1123.

Fang, Vivian W, Xuan Tian and Sheri Tice. 2013. "Does Stock Liquidity Enhance or Impede Firm Innovation?" *Journal of Finance, Forthcoming* .

Ferreira, Daniel, Gustavo Manso and André C Silva. 2012. "Incentives to innovate and the decision to go public or private." *Review of Financial Studies* p. hhs070.

Gopalakrishnan, V and Mohinder Parkash. 1995. "Borrower and lender perceptions of accounting information in corporate lending agreements." *Accounting Horizons* 9:13–13.

Gorton, Gary and Andrew Winton. 2003. "Financial intermediation." *Handbook of the Economics of Finance* 1:431–552.

Gottschalk, Peter and Robert Moffitt. 2009. "The Rising Instability of U.S. Earnings." *Journal of Economic Perspectives* 23(4):3–24.
URL: *http://www.aeaweb.org/articles?id=10.1257/jep.23.4.3*

Gourieroux, C., A. Monfort and E. Renault. 1993. "Indirect inference." *Journal of Applied Econometrics* 8(S1):S85–S118.
URL: *http://dx.doi.org/10.1002/jae.3950080507*

Guiso, Luigi, Luigi Pistaferri and Fabiano Schivardi. 2005. "Insurance within the Firm." *Journal of Political Economy* 113(5):1054–1087.
URL: *http://www.jstor.org/stable/10.1086/432136*

Hall, Bronwyn H, Adam B Jaffe and Manuel Trajtenberg. 2001. The NBER patent citation data file: Lessons, insights and methodological tools. Technical report National Bureau of Economic Research.

He, Jie Jack and Xuan Tian. 2013. "The dark side of analyst coverage: The case of innovation." *Journal of Financial Economics* 109(3):856–878.

Hsu, Po-Hsuan, Xuan Tian and Yan Xu. 2014. "Financial development and innovation: Cross-country evidence." *Journal of Financial Economics* 112(1):116–135.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru and Noah Stoffman. 2012. Technological innovation, resource allocation, and growth. Technical report National Bureau of Economic Research.

Kopczuk, Wojciech, Emmanuel Saez and Jae Song. 2010. "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937." *The Quarterly Journal of Economics* 125(1):91–128.
URL: *http://www.jstor.org/stable/40506278*

Lazear, Edward P. 1986. "Salaries and Piece Rates." *The Journal of Business* 59(3):405–431.
URL: *http://www.jstor.org/stable/2352711*

Lemieux, Thomas, W. Bentley Macleod and Daniel Parent. 2009. "Performance Pay and Wage Inequality." *The Quarterly Journal of Economics* 124(1):1–49.
URL: *http://www.jstor.org/stable/40506223*

Lemieux, Thomas, W. Bentley MacLeod and Daniel Parent. 2012*a*. "Contract Form, Wage Flexibility, and Employment." *American Economic Review* 102(3):526–31.
URL: *http://www.aeaweb.org/articles?id=10.1257/aer.102.3.526*

Lemieux, Thomas, W. Bentley MacLeod and Daniel Parent. 2012*b*. "Performance Pay and the Autocovariance Structure of Earnings and Hours.". Working paper.

MacLeod, W. Bentley and Daniel Parent. 1998. Job Characteristics and the Form of Compensation. CIRANO Working Papers 98s-08 CIRANO.
URL: *https://ideas.repec.org/p/cir/cirwor/98s-08.html*

Matzkin, Rosa L. 2007. Nonparametric identification. In *Handbook of Econometrics*, ed. J.J. Heckman and E.E. Leamer. Vol. 6 of *Handbook of Econometrics* Elsevier chapter 73.
URL: *https://ideas.repec.org/h/eee/ecochp/6b-73.html*

Meghir, Costas and Luigi Pistaferri. 2004. "Income Variance Dynamics and Heterogeneity." *Econometrica* 72(1):1–32.
URL: *http://dx.doi.org/10.1111/j.1468-0262.2004.00476.x*

Naidu, Suresh and Noam Yuchtman. 2016. Labor Market Institutions in the Gilded Age of American Economic History. NBER Working Papers 22117 National Bureau of Economic Research, Inc.
URL: *https://ideas.repec.org/p/nbr/nberwo/22117.html*

Naidu, Suresh, Yaw Nyarko and Shing-Yi Wang. 2016. "Monopsony Power in Migrant Labor Markets: Evidence from the United Arab Emirates." *Journal of Political Economy* 124(6):1735–1792.
URL: *https://ideas.repec.org/a/ucp/jpolec/doi10.1086-688877.html*

Nanda, Ramana and Matthew Rhodes-Kropf. 2013. "Investment cycles and startup innovation." *Journal of Financial Economics* 110(2):403–418.

Nini, Greg, David C Smith and Amir Sufi. 2012. "Creditor control rights, corporate governance, and firm value." *Review of Financial Studies* 25(6):1713–1761.

Piketty, Thomas and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913 - 1998." *The Quarterly Journal of Economics* 118(1):1–41.
URL: *+ http://dx.doi.org/10.1162/00335530360535135*

Piketty, Thomas and Emmanuel Saez. 2006. "The Evolution of Top Incomes: A Historical and International Perspective." *American Economic Review* 96(2):200–205.
URL: *http://www.aeaweb.org/articles?id=10.1257/000282806777212116*

Postel-Vinay, Fabien and HÃ©lÃšne Turon. 2010. "ON-THE-JOB SEARCH, PRODUCTIVITY SHOCKS, AND THE INDIVIDUAL EARNINGS PROCESS*." *International Economic Review* 51(3):599–629.
URL: *http://dx.doi.org/10.1111/j.1468-2354.2010.00594.x*

Roberts, Michael R and Amir Sufi. 2009. "Renegotiation of financial contracts: Evidence from private credit agreements." *Journal of Financial Economics* 93(2):159–184.

Rohwedder, Susann and Jeffrey B. Wenger. 2015. Worker Misclassification and the Hours and Earnings Effects of Expanded Coverage. Working Papers WR-1114 RAND Corporation.
URL: *https://www.rand.org/pubs/working_papers/WR1114.html*

Rosen, Sherwin. 1985. "Implicit Contracts: A Survey." *Journal of Economic Literature* 23(3):1144–1175.
URL: *http://www.jstor.org/stable/2725461*

Smith Jr, Clifford W and Jerold B Warner. 1979. "On financial contracting: An analysis of bond covenants." *Journal of financial economics* 7(2):117–161.

Solow, Robert M. 1957. "Technical change and the aggregate production function." *The review of Economics and Statistics* pp. 312–320.

Tian, Xuan and Tracy Yue Wang. 2014. "Tolerance for failure and corporate innovation." *Review of Financial Studies* 27(1):211–255.

Trejo, Stephen J. 1991. "The Effects of Overtime Pay Regulation on Worker Compensation." *The American Economic Review* 81(4):719–740.
URL: *http://www.jstor.org/stable/2006639*

Trejo, Stephen J. 1993. "Overtime Pay, Overtime Hours, and Labor Unions." *Journal of Labor Economics* 11(2):253–278.
URL: *http://www.jstor.org/stable/2535281*

Trejo, Stephen J. 2003. "Does the Statutory Overtime Premium Discourage Long Workweeks?" *Industrial and Labor Relations Review* 56(3):530–551.
URL: *http://www.jstor.org/stable/3590923*